P.D. 1993
p. 308-430 (123)
XP-000860006
H04Q7/22
97374-EPD1-D1
3/13/1

Radio Resource Management

6

RADIO RESOURCE MANAGEMENT

As a telecommunication system, GSM enables its users to communicate through transmission paths of various characteristics, as explained in the previous chapters. However these transmission paths are not reserved once for all between any two pairs of users. They are set up on demand, and only for the time necessary for a given communication. This requires exchanges of information, not only between the users and the network they are in direct contact with, but also between machines within the network. This and the two next chapters are devoted to the description of these information exchanges, which enable distant participants, users and machines, to act together to provide the communication services for which the networks are designed.

This technological field is known as signalling, and under this name, its reputation has not yet grown as wide as other fields such as modulation, signal processing, and other transmission techniques. Many people consider it simply as a branch of software engineering, though it is in fact at the centre of the design of complex systems, where tasks can be executed only through the co-operation of distinct machines. This is the case of telecommunications, where machines are by essence distinct and distant. The study of GSM signalling is therefore of prime importance to understand how the system operates, and the reader should not be surprised that half of the book is devoted to the signalling interchanges.

Signalling is often the juxtaposition of many simple and more or less inter-dependent procedures, and its complexity stems mainly from the number and the diversity of small issues. We have already seen in

Chapter 2, concerning the architecture, that the basic methodology to tackle such issues is "divide and conquer". Pervasive in the specifications of GSM signalling is a split in three functional domains: Radio Resource Management, Mobility Management and Communication Management. The management of the calls is the upper plane in this organisation, and deals with the establishment and the release of end-to-end transmission paths, through the GSM domain and by interworking with external networks, to support user to user communications. Communication Management, as a functional plane, relies on the Mobility Management functions for dealing with the mobility of the users and with security-related functions. The management of the radio resources groups functions specific to the radio interface.

A major difference between a radio mobile telecommunication network and a network with fixed links is the management of the access resources. In a fixed system, a dedicated communication medium exists continuously between the user terminal and the infrastructure, ready to be used when a call needs to be established. On the contrary, in a cellular system like GSM, a dedicated channel over the radio interface is provided to the mobile stations only on demand and for the duration of the call, under the control of the infrastructure. This calls for functions which bear no equivalent in ISDN for instance. Even if 64 kbit/s channels are allocated dynamically in the case of an ISDN multi-terminal installation, this resource management is rather limited compared to the one in a full GSM cell. Moreover, in ISDN, a signalling channel is always ready for use by any terminal. The matter is quite different in GSM, where the signalling capabilities offered to a mobile station in idle mode, that is to say when not allocated a radio channel to its private usage, are limited to the absolute minimum. The consequence is that a host of new procedures are needed.

Besides dynamic channel allocation, another feature of GSM (or cellular systems in general) compared with fixed networks is the **handover**. The problem consists in providing a dedicated channel from mobile station to MSC at every moment during a call, despite the movements of the user. This calls for a complex measurement and decision process to trigger the transfer of the communication at the right moment and toward the right cell. In a cellular system, the handover process is very important, since it impacts significantly the quality of the communications as perceived by the users, as well as the spectral efficiency.

This chapter will be devoted to these topics, that is to say to the functions required to co-ordinate the mobile stations and the infrastructure so as to provide the suitable transmission means over the radio interface, whatever the telecommunication service requires, and

whatever the user's movements. These functions form a well defined area, which we presented in Chapter 2 as the **RR** (Radio Resource management) functional plane. They are spread among four entities in the canonical GSM architecture: the mobile station (of course!), the two base station sub-system components (BTS and BSC), and a small part in the MSC. All the higher layer functions, described in Chapters 7 and 8, are basically managed directly between mobile station and MSC, the base station sub-system (BSS) acting for these functions as a single complex transmission system. The spread of the radio resource management functions implies the existence of signalling procedures between the involved infrastructure machines; this is the purpose of the signalling protocols on the A (BSC-MSC) and Abis (BSC-BTS) interfaces, which will therefore be described in this chapter.

The chapter is basically composed of two parts. After preliminary architecture considerations needed to introduce some specific concepts, the major requirements which drove the design of the RR protocols will be looked at. Then, after a section to present the protocol architecture, the various procedures needed to fulfil these requirements will be developed.

Preliminary Architecture Considerations

A section in the middle of this chapter will be devoted to the architecture and the protocols in the Radio Resource management domain. However, some basic notions are necessary for the understanding of the first part of the chapter. It concerns mainly the notion of anchor MSC.

The major roles are played in this chapter by the components of the BSS, the BTS and the BSC. The MSC intervenes a little, to deal with handovers between cells managed by different BSCs. Some handovers may even transfer the mobile station from a cell within one MSC area to a cell in another MSC area, thus involving *two* MSCs. The roles of the two MSCs are different. In no case does the MSC in charge of the communication relinquish its control to the new MSC. This MSC is called the **anchor MSC** for the connection. This is an important design choice of GSM, with numerous consequences on the procedures. Several arguments justify this choice; a compelling one is the charging problem, since toll ticketing is much simpler when one MSC follows a call from its beginning to its end.

A consequence of this architectural choice is that after an inter-MSC handover, two MSCs (and at most two) may be involved in the connection. The transmission chain between the mobile station and the interworking point with external networks is then composed of a BTS, a

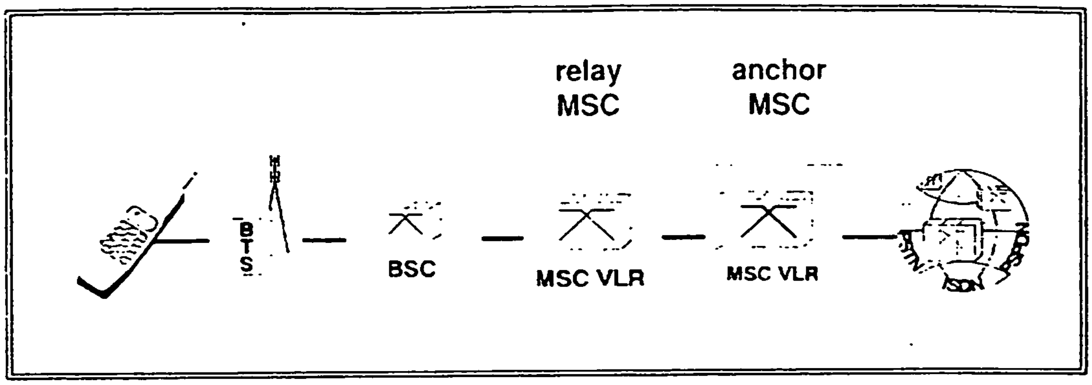


Figure 6.1 – The concepts of anchor and relay MSCs

The transmission chain may involve two (and at most two) MSCs:
the anchor MSC in charge of the communication management and
the relay MSC in charge of the BSS with which the mobile station is in contact.

BSC and either two MSCs, a **relay** MSC and the anchor MSC, or of a single MSC (see Figure 6.1). To ease the problem of terminology, a convenient approach is to consider the notions of relay MSC and of anchor MSC as functional, and to admit that when there is only one MSC, it is at the same time the relay MSC and the anchor MSC. Thus the term relay MSC will be used to refer to the MSC in direct contact with the BSC, even if it is also the anchor MSC; and the term anchor MSC will be used to refer to the MSC in charge of upper layer treatments, even if it is also the relay MSC.

6.1. RR FUNCTIONS

In this first half of this long chapter, we will study the different Radio Resource management aspects from the requirement side. While implementation issues will often be addressed, in particular for the distribution of the tasks between the involved machines, the details of the signalling procedure will be entirely in the second part.

The functions covered by the management of radio resources are centred on the management of transmission paths over the radio interface, and more exactly between the mobile station and the anchor MSC. To develop these functions, we will use the concept of RR-session, which will be presented first. After a small passage concerning the access and the paging, through which things start, we will deal with the handling of

the main properties of the transmission chain, such as whether signalling, speech or data is transported, and whether ciphering is applied or not.

The next issue will be on how handovers are decided. The handover execution itself will be treated mainly in the procedural section. Addressing the handover preparation issue will take us deep into considerations about the measurements performed by the mobile station and the base station. These measurements are the basic information upon which handovers can be decided.

Next, two ancillary functions of the transmission over the radio interface will be looked at, the management of the transmission power and of the timing advance.

Finally we will deal with the management of the radio channels on the radio interface as a whole set. The two main facets are the handling of the configuration of the radio channels in each cell, and the allocation strategy of the dedicated radio channels (TACH/8s and TACH/Fs).

6.1.1. THE CONCEPT OF RR-SESSION

Most of the functions in the Radio Resource management plane relate to the management of the transmission between the mobile station and the anchor MSC. For each mobile station engaged in a communication, there exists a transmission path, as well as a signalling path, between itself and the anchor MSC. As seen from a mobile station, such a path is set up when it enters the dedicated mode (i.e., when it leaves the idle mode), and is released when the mobile station goes back to idle mode. In the infrastructure, a transmission path exists for all this period, but can be thoroughly modified, especially by handovers. We will refer to what is managed during this period of time as an **RR-session**. As a minimum, an RR-session must include means to transmit signalling between the mobile station and the anchor MSC through the BTS, BSC and possibly a relay MSC, including a dedicated radio channel, references to manage it on both the BTS-BSC interface and the BSC-MSC interface, and means in the BSS to monitor the radio connection and take handover decisions when necessary. This minimum set suffices only in the cases where the transfer of circuit-type user data is not required, such as for location updating, short message transfer or supplementary service management. When circuit-type user data needs to

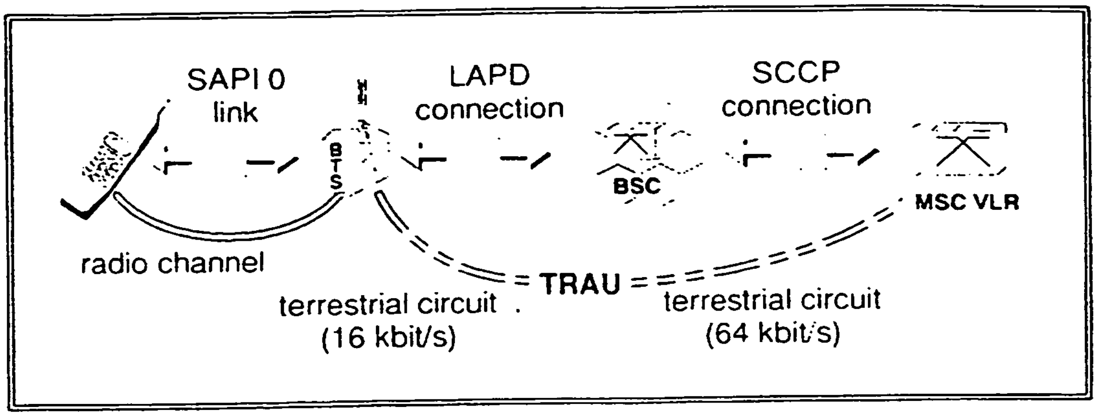


Figure 6.2 – Contents of an RR-session

An RR-session contains both the signalling resources between mobile station and anchor MSC, including a dedicated channel on the radio path, as well as the user data circuits if need be.

be transmitted, then a complete circuit connection between mobile station and anchor MSC is also part of the RR-session, as shown in figure 6.2. For instance, a speech call requires the use of a signalling connection as well as a speech-carrying connection between mobile station and MSC. This last connection makes use of dedicated resources such as the speech transcoder transforming the GSM-specific speech representation into the 64 kbit/s representation used in fixed networks.

An RR-session has many different characteristics which have to be managed by procedural means. First, two different kinds of dedicated channels exist on the radio interface; they have been referred to as TACH/8 and TACH/F in Chapter 4 (there will be three when "half-rate" channels are included: TACH/8, TACH/H and TACH/F). Second, when a circuit for user data is present, it can be used according to different transmission modes. Finally, some other less important transmission peculiarities characterise RR-sessions. An example is whether ciphering is applied or not. All these characteristics may change during the lifetime of an RR-session.

In the *Specifications*, the term "RR-connection" refers to what is managed during the period of connection to a given BSC. A change of BSC (e.g., at inter-BSC handover) entails a change of the RR-connection. The *Specifications* do not have a specific term covering what is managed for the whole period where the mobile station is in dedicated mode between two periods of idle mode, i.e., what we call here an RR-session. Figure 6.3 illustrates the concepts of RR-connection and RR-session, and their relationship. The figure also shows that an RR-session can be used for several calls in succession or in parallel, or more generally several CM-transactions (CM for Communication Management), as will be

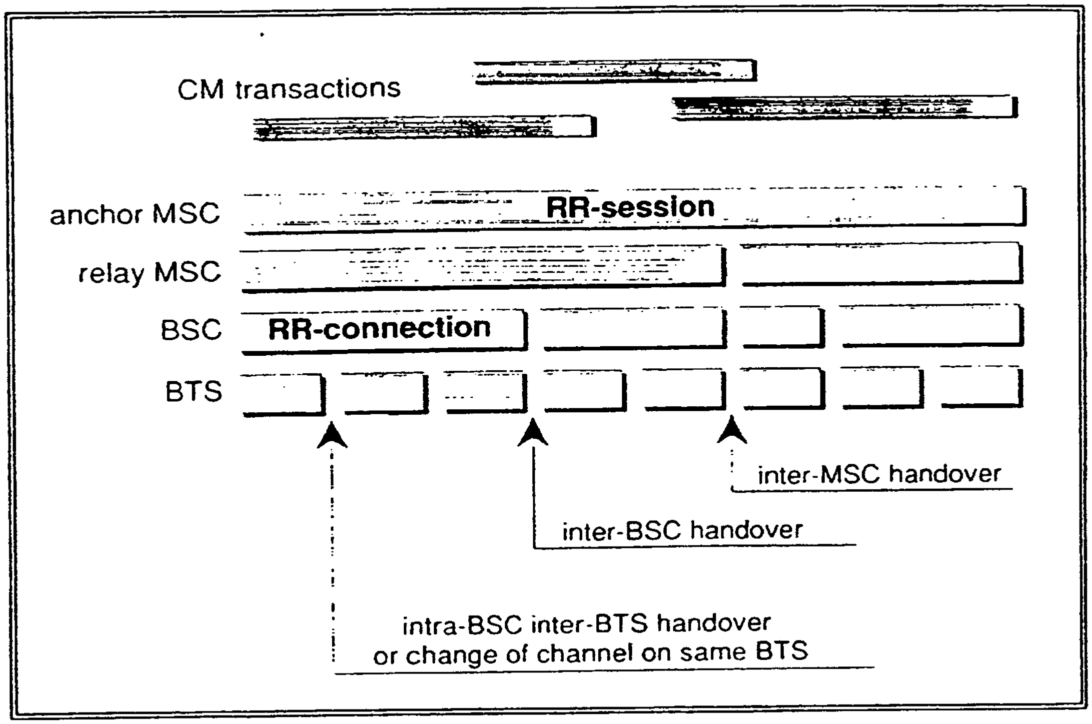


Figure 6.3 – The concepts of RR-session and RR-connection

From the concept of radio connection (bottom line) to the one of RR-session (top line), different levels of transition awareness may be defined.
The *Specifications* use the concept of RR-connection, which corresponds to the BSC view.
CM-transactions may run in parallel or in tandem during the lifetime of an RR-session, as shown on the top of the figure.

described in Chapter 8. The beginnings and ends of CM-transactions relate to the usage of the transmission, and are completely independent from RR-connections, whose succession relates to the movement of the mobile stations.

The RR-session is the bond between the two domains of radio resource management and communication management. It represents the views of the mobile station and of the anchor MSC. An RR-session starts when the mobile station goes to dedicated mode (the access, when the initial assignment of dedicated channels is performed), and disappears when the mobile station goes back to idle mode.

The life of an RR-connection is punctuated by intra-BSC handovers and changes of radio channels, and this defines another subdivision in the lifetime of a RR-session. At the lower level, the channel connection corresponds to the continuous usage of the same radio channel by the same mobile station. A channel connection starts either through an initial assignment, a subsequent assignment (a change of channel done, e.g., because the allocated channel is no more of the

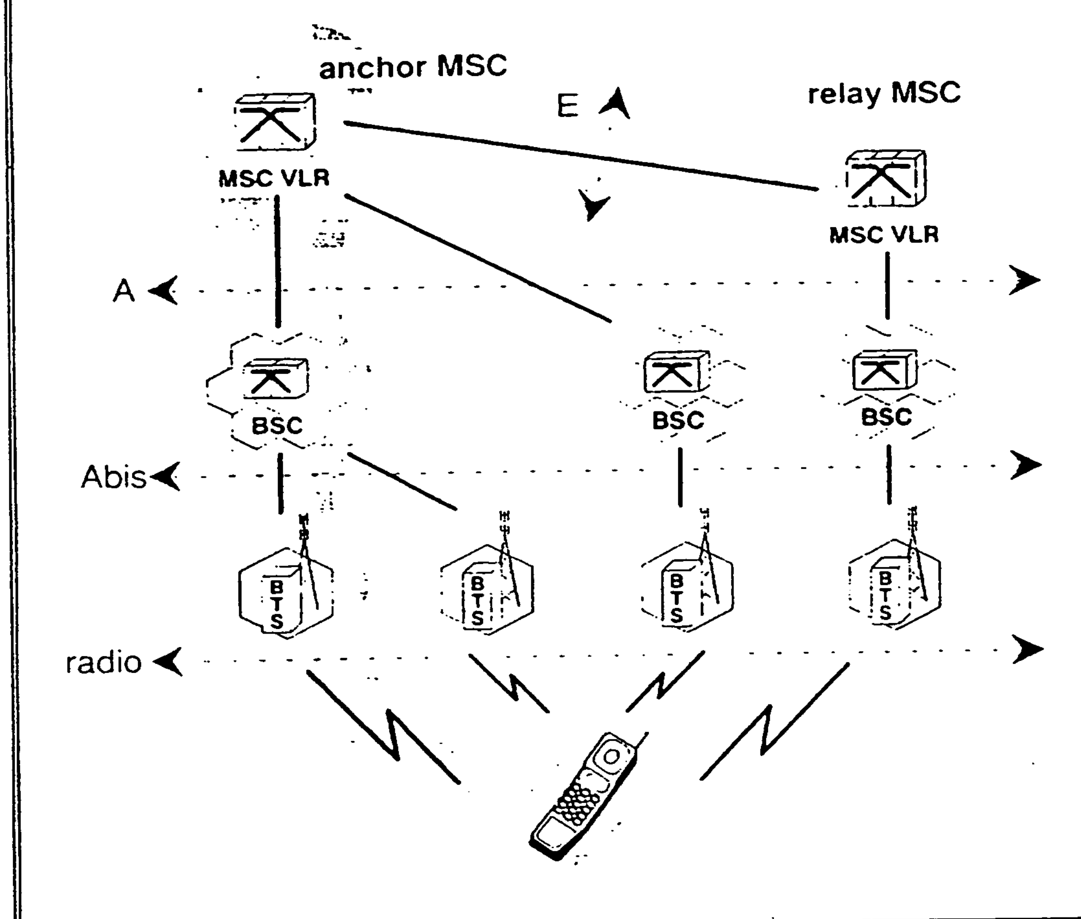


Figure 6.4 – Configurations changes for an RR-session

During its lifetime, an RR-session may go through different transmission configurations. If the initial configuration is for example the left one, it may be changed to any one of the others shown. The anchor MSC remains in charge of the upper layers.

needed type) or an incoming handover of any kind. It disappears either through the release of the RR-session, the assignment of another channel or an outgoing handover. Channel connections represent the view of BTSs. By design choice, a BTS considers a change of radio channel within the same cell as two independent channel connections. When a channel connection is cleared, the related data is wiped out in the BTS, regardless of whether the mobile station is allocated a new channel of the same BTS or in another one.

There are very few stable characteristics of the RR-session beside the corresponding contexts in the mobile station and in the anchor MSC, especially when one recalls that the physical path of the transmission may change thoroughly when a handover occurs. At a given moment in time, a given RR-session is managed by one BTS, one BSC, an anchor MSC and

sometimes in addition a relay MSC, as shown in figure 6.4. Each of these machines maintains some context related to the RR-session. When a handover occurs, the configuration changes, some contexts must be erased and others must be created in other machines. Based on the corresponding configurations, the functional architecture of GSM distinguishes three kinds of handover. In an intra-BSC handover, only the radio channel, the context in the BTS and possibly the BTS are changed. In an (intra-MSC) inter-BSC handover, the BSC is changed in addition to the radio channel and the BTS. Finally, in an inter-MSC handover, the relay MSC is either created, replaced or suppressed. In all cases, the anchor MSC remains in place throughout all the life of the RR-session. It is indeed the only machine sure to be a constant, and the context of the RR-session in the anchor MSC is indeed the anchorage point of the session.

6.1.2. INITIALISATION

A mobile station has two widely different operating modes, the idle mode, when it is not engaged in a connection with the infrastructure, and the dedicated mode, when a full duplex channel enables actual communications to take place. The transition from idle to dedicated mode is the first step of the initialisation of an RR-session, and is called the **access**. As part of an initialisation process, it has many functional aspects, and this section will address only some aspects specific to the access issue. The full-blown detailed procedure itself will be presented much later in the chapter, after all the relevant concepts will have been introduced.

Access may be triggered either to fulfil a need expressed first on the mobile station side (e.g., a call originated by the user of the mobile station, but also a location updating), or on the infrastructure side (e.g., a call toward the GSM subscriber). In all cases the access procedure is the same, and is initiated by the mobile station. When the network desires the establishment of an RR-session, it **pages** the mobile station which constitutes a request for it to access. To this avail, when normal service is provided, the mobile station listens in idle mode to a paging sub-channel, part of the PAGCH. If a message on this sub-channel indicates that its subscriber is paged, the mobile station starts the access procedure, as it does when the user requests it. We will then deal separately with the two aspects, the access proper, and the paging.

6.1.2.1. Access

The mobile station manifests its will to access by sending a message on the random access channel (RACH), which is answered by an

initial assignment message on the paging and access grant channel (PAGCH), carrying the description of the allocated dedicated channel. The mobile station provides very little information in its access request (the message has only 8 bits of contents). In particular the mobile station does not give its identity at this moment, nor the detailed reason for the access. Another particularity of interest is that the access on the RACH is not regulated. The consequence is that two mobile stations may send access requests simultaneously, which would result most often in neither being received by the BTS. Most of the complexity of the access procedure (repetition of the attempt, resolution of access to the same channel by two mobile stations) comes from fixing these problems.

The access ends with the allocation of a radio channel for the use of the mobile station. This is called the **initial channel assignment**, referred to in the *Specifications* as immediate assignment. The access procedure, though limited to the means needed for the transition between the two modes, makes exclusive use of two specific channels, the PAGCH and the RACH.

From a more general point of view, the access function is the initiation of an RR session. As such it includes the initialisation of all the contexts and all the recurrent functions which are part of the RR session. The access will therefore be revisited in many of the functional sections, such as those dealing with the management of the timing advance, with the transmission power control (which must be started during the access), and with the channel allocation. The detailed description of the access procedure, in the last part of this chapter, will provide the synthesis of these different aspects.

6.1.2.2. Paging and Discontinuous Reception

Compared to the other functions described in this chapter, paging is somewhat particular, because it is not directly linked to an RR-session. What is indeed the relationship between paging and radio resource management? Since RR-sessions are only established at the initiative of the mobile station, the network infrastructure needs some means to trigger such an establishment; this role is indeed fulfilled by the paging procedure. But there is no common reference, no clear relationship between the paging message and the ensuing RR-session establishment. The only clue for the network is the indication by the mobile station in the first dedicated message that the RR-session has been established in response to some unspecified paging indication.

Paging is in some sense closer to Mobility Management functions than to Radio Resources management functions, as it serves to locate a

mobile station within a whole location area. The grouping of the paging function with RR management, which is also the one followed in the *Specifications*, reflects the relationship which exist between paging and a number of true RR functions. For instance, paging messages and initial assignment messages share the same channel (the PAGCH). This approach is also sensible from a pragmatic point of view since the main job for paging is done by the BSS, which is otherwise only concerned with RR functions.

How does paging start? When an incoming call arrives, the MSC/VLR requests the BSS to perform a paging in some of the cells of the BSS. The MSC provides the concerned BSC(s) with the identity of the mobile subscriber to be paged and the list of cells in which the paging should be issued. The BSC is in charge of managing the PAGCH, i.e., the grouping and scheduling of paging messages as well as initial assignment messages. This scheduling may be more or less optimised, depending on the manufacturer. The *Specifications* describe a framework for such an operation, but the operator/manufacturer can choose how often to repeat unanswered paging messages, whether to send initial assignment messages also on those parts of the PAGCH which correspond to the paging sub-channels, etc.

The split of functions between BTS and BSC with regard to paging also allows some flexibility and is somewhat manufacturer-dependent. Typically, the high-level tasks, such as priority decisions, rest with the BSC. Sophisticated approaches may take the system load into account for the respective priority of paging and initial assignment messages.

Another aspect of paging also tackled by the BSS concerns the concept of discontinuous reception. For the sake of battery consumption in handheld mobile stations, it is important to minimise the amount of information the mobile station has to receive, demodulate and analyse while in idle mode. To this avail, the downlink common control channel can be divided into several paging sub-channels, and all paging messages pertaining to a given subscriber are normally sent on the same sub-channel. The sub-channel organisation can vary on a cell basis, but broadcast information allows the mobile stations to determine it. Such a scheme allows mobile stations to restrict their monitoring of paging messages to their own paging sub-channel, thereby increasing significantly the lifetime of their battery, at the expense of a small increase in delay for the setting up of incoming calls. Such a feature is referred to as discontinuous reception, or DRX, and is not to be confused with discontinuous transmission (DTX) with which it bears no relationship except the similar names. Mobile subscribers are assigned to paging sub-channels in a pre-determined way taking into account the three last digits of their international mobile subscriber identity (the

IMSI), so that the knowledge of the PAGCH configuration is enough for each mobile station to determine which blocks of which CCCH carrier unit it should listen to.

The PAGCH follows a 51×8 BP cycle, where BP denotes a burst period, using 9 blocks per cycle for a PAGCH/F (the PAGCH of large capacity) and 3 blocks per cycle for a PAGCH/T (the PAGCH of small capacity), as described in Chapter 4. A certain (parameter-controlled) number of these blocks belong to some paging sub-channel, the others being reserved to initial assignment messages. This number may range from 2 to 9 for a PAGCH/F and from 1 to 3 for a PAGCH/T. A paging sub-channel is (almost) cyclic, with a cycle ranging from 2 to 9 times 51×8 BP (that is to say from 0.95 second to 4.25 seconds), here again under the control of a parameter. Hence, on a given PAGCH/F, there can be from 4 to 81 paging sub-channels (2 to 27 for a PAGCH/T). The two parameters describe the PAGCH configuration and are broadcast on the BCCH. The choice of these parameters is the operator's. The second parameter (linked to the cycle of paging sub-channels) corresponds to a compromise between the access time and the power consumption of the mobile stations. The first one was introduced solely to enable the development of very simple PAGCH scheduling algorithms which do not use the possibility to send initial assignment indications on a paging sub-channel. In such cases, the choice of the parameter is related to the ratio between the paging load and the initial assignment load. Otherwise, the parameter is set so that any PAGCH block belongs to some paging sub-channel.

A small detail is that the interval between two successive paging sub-blocks of the same sub-group is constant, except (in some combinations) once every 3.5 hours, when the numbering scheme goes through 0. This happens when the paging sub-channels "cycle" is not a divider of the numbering cycle (lasting $2048 \times 26 \times 51 \times 8$ BP), i.e., for cycles of 3, 5, 6, 7 or 9 times 51×8 BP.

Procedural Requirements for Paging

The procedural requirements for paging management include means for the MSC to require a given subscriber to be paged, a mechanism for the BSC to control the sending of this paging message (or alternatively to provide the BTS with the relevant data to build and schedule the paging messages) and a way to indicate the PAGCH configuration to all mobile stations. In addition, some means to configure the PAGCH are required in the operation sub-system (OSS), as part of the more general techniques for controlling the cell configuration. This is dealt with in Chapter 9.

6.1.3. TRANSMISSION MANAGEMENT

The life and deeds of the RR-sessions have only been sketched in this previous section. We will now see in more detail what constitutes the management of the transmission characteristics of an RR-session. These characteristics are chosen depending on the service to be provided. They are decided by the anchor MSC, but transitions are co-ordinated by the BSC. The most obvious area of co-ordination relates to the type of data which must be carried. On this depends the existence or not of terrestrial circuits, the type of the radio channel and the transmission mode (speech, or data with some data rate). Beside these, the BSS has also to manage the cipher mode, as well as the discontinuous transmission modes.

We will take each of these aspects in turn, to examine what is to be done.

6.1.3.1. Transmission Mode Management

We group under the term **transmission mode** the main transmission characteristics. More precisely, the concept of transmission mode refers to the way the GSM transmission chain is used for carrying circuit user information, from the mobile station to the point of interconnection with partner networks. The set of possible transmission modes differs depending on the type of channel used on the radio interface. The transmission modes have already been described in detail in terms of their "physical layer" characteristics in Chapters 3 and 4. Table 6.1 summarises which transmission modes exist on each radio

TACH/8	TACH/F	*TACH/H*
signalling only	signalling only speech data 3.6 kbit/s data 6 kbit/s data 12 kbit/s, transparent data 12 kbit/s, non-transparent	*signalling only* *speech* *data 3.6 kbit/s* *data 6 kbit/s, transparent* *data 6 kbit/s, non-transparent*

Table 6.1 – Transmission modes used on the radio channels

The only transmission mode on a TACH/8 corresponds to signalling only,
whereas the "full-rate" channels may carry 6 different modes,
and the "half-rate" channels will presumably carry the 5 modes listed.
The "transparent" and "non-transparent" modes (see Chapter 3)
use the same channel coding on the radio path,
but lead to variations in the transmission functions within the BSS.

channel type. The transmission modes which were defined on half-rate channels are not in the phase 1 *Specifications*, but are also shown in this table.

The "signalling only" mode corresponds to the non-usage of the channel to carry circuit-type user data. The transport of signalling information is a capability which exists in all transmission modes, once an RR-session has been established. There are even cases where it represents the only transmission need; for instance, at the beginning of a call (before the conversation phase), for transmission of user data other than circuit-type (short message transmission), for location updating or for supplementary service management.

The existence of a full transmission path including terrestrial circuits between the BTS and the MSC/IWF, and the inclusion in this path of a transcoder and rate adaptor unit (TRAU), depend on the mode; for instance, there are none of these in "signalling only" mode.

In general, the transmission mode is chosen by the MSC, depending on the end-to-end service. When the RR-session is initially established, the MSC does not intervene in the process before the point when it knows exactly the transmission needs to fulfil; up to this point, the *Specifications* impose the mode "signalling only". The channel can be chosen (by the BSC) to be of any type, though it is typically a TACH/8 (see page 355 for the allocation strategies which can be followed by the BSC). Later during the lifetime of the RR-session, the channel type and the transmission mode may change; these changes are decided by the MSC, in order to adapt the transmission media to the needs of the users.

Although the decision lies with the MSC, it is the BSC which chooses the exact channel (of the requested type) and is in charge of co-ordinating the different machines, including the mobile station. The only exception to this rule is the establishment of the terrestrial circuit (between BSC—or TRAU—and MSC), which is always done by the MSCs. This exception stems from no specific reason but the weight of history. Switches have always been in charge of establishing their surrounding circuits. This situation actually raises some problems in phase 2, and it would probably have been a better choice to let the BSC in total control of the whole transmission chain. The management of the terrestrial circuit will be addressed in a bit more detail in the next section.

Procedural Needs for Transmission Mode Management

The procedural organisation of transmission mode management has two facets. First the MSC must be able to indicate the need for a

change in the transmission mode at any moment during an RR-session. Second, the BSC must have means to co-ordinate the mobile station, the BTS and the TRAU for fulfilling the MSC request. This last aspect is split into two cases as seen from the BSC and the mobile station, whether the type (full rate or eighth rate) of the existing radio channel fits the requirements or not. A typical case when it does not is at the beginning of an RR-session established for the purpose of establishing a call, if the existing channel is a TACH/8. In such a case, the radio channel must be changed. The procedure to change the radio channel used by an RR-session without changing the cell is called a **subsequent assignment**. If the type of the existing channel is appropriate, but the transmission mode is not, the procedure between the mobile station and the BSC is a **mode modification**.

It should be noted that the handover procedure can be used for changing the transmission mode, including the type of channel. Reciprocally, an intra-cell change of channel is often called a handover, if triggered by quality considerations and not by the adaptation to the needs of the service. A small digression on the terms is useful at this stage. This last usage of the term handover as to be understood from the point of view of *why* it is done. However, from the point of view of *what* is done as seen between the BSC and the mobile station, there is indeed no difference between a subsequent assignment and an intra-cell "handover". The accepted use of the term handover thus depends on the context.

6.1.3.2. Terrestrial Channel Management

An RR-session may or may not include a full circuit between the mobile station and the MSC to carry user information. This circuit is not always present; for instance, if the RR-connection is used for location updating, such a circuit is of no use.

When it is present, it includes a radio channel (a TACH/F, or in the future a TACH/H), and terrestrial circuits connecting the BTS with the anchor MSC via the BSC and a TRAU if applicable. The terrestrial circuit from BTS to BSC (possibly via the TRAU) is in a one-to-one relationship with the radio channel, and is then dealt with by the BSC as part of the radio channel management.

On the A interface, the circuit (from the BSC or TRAU to the MSC) is allocated to an RR-session by the relay MSC. The indication that a circuit has been allocated is given to the BSC through signalling

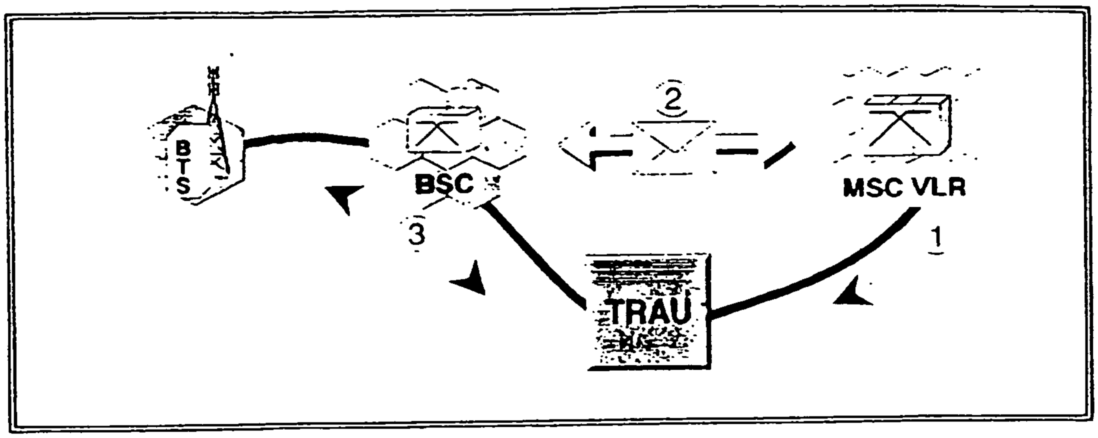


Figure 6.5 – Channel allocation on the terrestrial interfaces

When the TRAU is situated on the MSC-side of the BSC, the MSC first chooses a circuit towards a TRAU, then a signalling exchange takes place on the A interface, and finally the BSC controls the set-up of circuits between BTS and TRAU.

means, enabling the BSC to then connect it to the radio part of the path through its switching matrix (see figure 6.5).

Finally, the circuit between the anchor MSC and the relay MSC, if they are different, is established in the canonical architecture using standard PSTN or ISDN methods. The establishment is initiated by the anchor MSC, using the same procedural means it would use to establish a call.

The existence of a remote transcoder and rate adaptation unit along the BTS to relay MSC path makes this picture somewhat more complex. If no BSC-controlled switching matrix exists between the TRAU and the MSC, then a one-to-one relationship should exist between a specific transmission resource in the TRAU and an MSC-TRAU circuit. Thus, it is in that case the MSC which implicitly chooses the transcoding device by choosing the terrestrial circuit. This situation bears no consequence if the TRAU devices are all equivalent; otherwise, a potential problem exists, since it is the BSS which is in charge of the TRAU, and of the consistency between what the TRAU does and what is done on the radio interface. Thus, the TRAU is in some way managed both by the MSC and the BSS, and in practice, the architectural choices of GSM make it difficult to have distinct types of TRAU.

Procedural Needs for Terrestrial Channel Management

The terrestrial part of the transmission path of an RR-session is established when a TCH/F is requested by the anchor MSC. This is a part of the subsequent assignment procedure. It is obviously modified at each

handover, and is in particular entirely changed for an inter-MSC handover. The establishment of the new path, and the release of the previous one are part of the handover execution procedure.

6.1.3.3. Cipher Mode Management

The transmission over the radio path has a few characteristics which must be managed independently from the type of transported data. The first one is the cipher mode: the transmission may be ciphered or not, as desired by the MSC, according to some criteria which depend on the operator.

An RR-session is always started in "signalling only" mode, and always in clear text (i.e., not enciphered). The latter is a necessary requirement, since the RR-session setup is performed without the network knowing the subscriber identity, and therefore ciphering with a user-related key cannot be applied. Means for transition from clear text to ciphered transmission on an existing RR-session are therefore required. The *Specifications* do not provide explicitly for the reverse transition, i.e., from ciphered mode to non-ciphered mode. No need was identified for such a transition. Similarly, the *Specifications* do not cater for the change of ciphering key while in ciphered mode. However, the existing procedures could be used in the future to support these transitions, should the need arise.

Similarly to the transmission mode management case, the transition from clear text transmission to ciphered mode is decided upon by the MSC, the BSC being in charge of co-ordinating the actual change. Cipher mode management impacts the mobile station and BTS.

Procedural needs include the means for the MSC to provide the ciphering parameters (the mode, and the user ciphering key *Kc* if needed) to the BSC for an incoming handover, the means for the MSC to require a change of the current mode from non-ciphered to ciphered, and the means for the BSC to co-ordinate the transition in the BTS and the mobile station. This last aspect is of primary importance, since messages sent in one mode will not be understood by the peer entity if this entity is set in the other mode, resulting in an unrecoverable loss of connection. The way to synchronise mobile station and BTS will be described in the procedural section.

6.1.3.4. Discontinuous Transmission

Discontinuous transmission (DTX) has been described in Chapters 3 and 4. When discontinuous transmission is applied, actual transmission

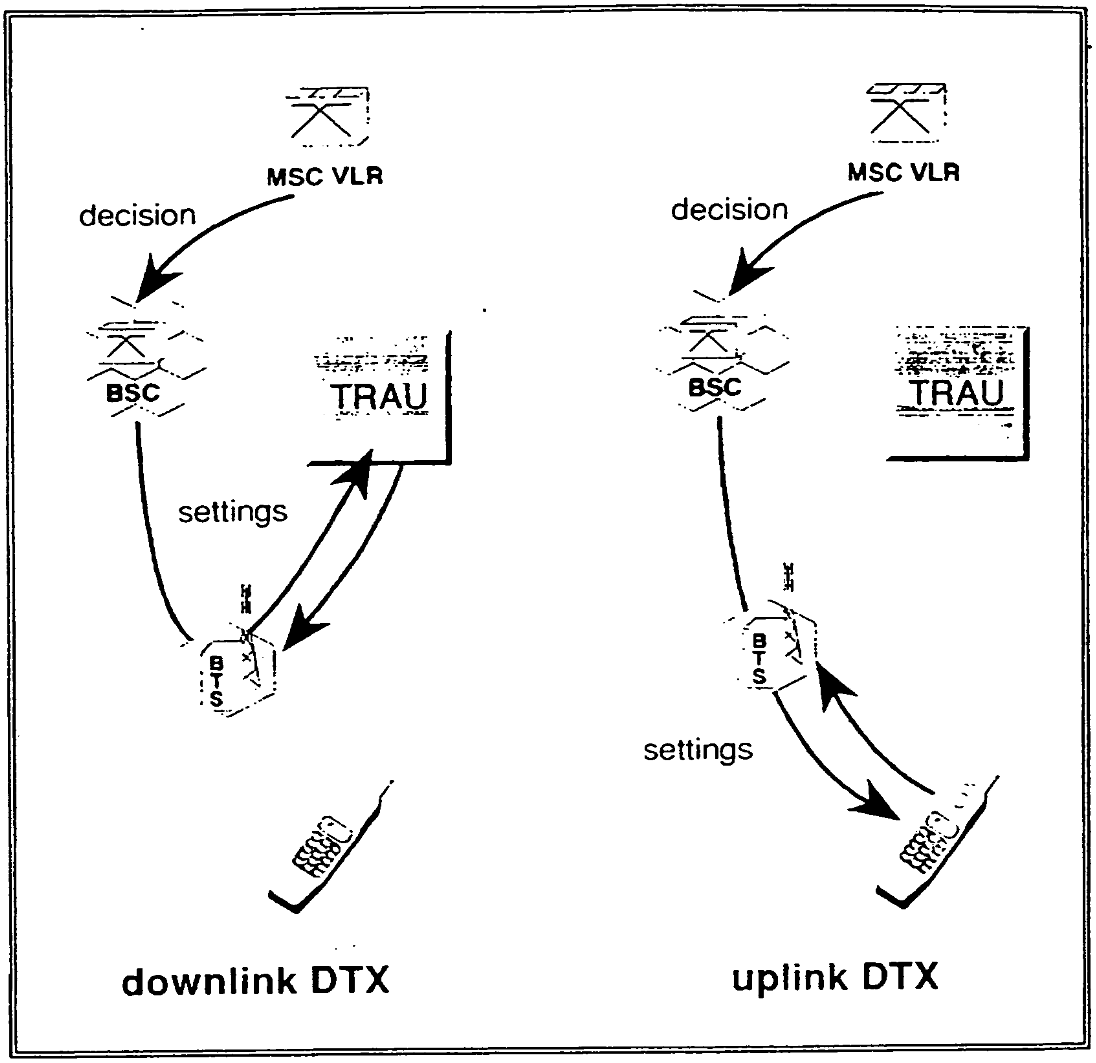


Figure 6.6 – Procedural needs for discontinuous transmission management

The use or not of discontinuous transmission needs only be notified to the transmitting end; the receiving unit does not need to know beforehand.

on the radio path is reduced to a minimum when it is detected that the user data does not contain meaningful information (during speech silences for instance). This feature is optional and must therefore be managed. Moreover, discontinuous transmission may be applied independently to each direction, so that the control of discontinuous transmission must take into account two components: the uplink mode and the downlink mode.

Discontinuous transmission is only relevant to some of the transmission modes, speech and non-transparent data, simply because in the other cases it is difficult to assess when user data transmission can be suspended without degrading the service. The discontinuous transmission mode affects the operation of the mobile station and of the TRAU. The

BTS is obviously concerned, but derives its behaviour dynamically from data coming from the mobile station (uplink) and from the TRAU (where it exists) and MSC/IWF (downlink). Whether discontinuous transmission should be applied or not is, there again, decided upon by the MSC, and the execution is controlled by the BSC. A GSM BSS must indeed be able to cope with discontinuous transmission, whatever the strategy of the corresponding MSC.

The choice of the strategy for applying discontinuous transmission is one of the many configuration parameters on which operators may play to optimise their network. Several considerations must be taken into account in this strategy. For instance, GSM mobile to mobile calls suffer a loss in quality when discontinuous transmission is applied to both radio segments; experts refer to this phenomenon as "double clipping". The operator may therefore well choose not to apply downlink discontinuous transmission for such MS-to-MS calls, if they can be identified.

The downlink discontinuous transmission mode can be changed when the transmission mode changes; these are indeed the only instants at which the *Specifications* allow a change in the discontinuous transmission mode, since no other needs were identified.

For the uplink discontinuous transmission mode, the network can at any moment either force the mobile stations in communication to use it, forbid them to do so or leave the choice open.

Procedural needs for discontinuous transmission management include the means for the MSC to indicate whether discontinuous transmission should be applied for uplink and for downlink, and means for the BSC to configure the mobile station and the TRAU via the BTS (see figure 6.6).

6.1.4. HANDOVER PREPARATION

The possibility to change the cell during an RR-session is a very important function in a cellular system, and the major source of complexity in the Radio Resource management plane. We have already touched upon some of the aspects related to the *execution* of a handover, i.e., what has to be modified when a handover occurs. However, this represents but the tip of the iceberg. The process which precedes it, the handover preparation, may be thought of as a "behind the scenes" activity, yet it is the most important topic, which conditions both spectral efficiency and the quality of service as perceived by the users.

The decision to trigger a handover, and the choice of target cell, are based on a number of parameters, and various reasons may trigger this

decision. These reasons will be studied first, then a description of the parameters affecting the choice will follow. Among these parameters, radio measurements performed by the mobile station and by the BTS are of foremost importance, and will be looked at in detail. The functional distribution of functions will be addressed last.

6.1.4.1. Handover Purposes

At first sight, the aim of handovers is to avoid loosing a call in progress when the mobile station leaves the radio coverage area of the cell in charge. Such cut-offs are very badly perceived by the users, and have an important weight in the overall perception of the quality of service by the subscribers. We shall call this type of handover "**rescue handover**", where a high probability exists that the call will be lost if the cell is not changed. An extreme form of the rescue handover is the call re-establishment, which is an attempt for the mobile station to salvage the connection *after* an effective loss of communication with the serving cell.

In other cases, it may be of interest to change the serving cell of a given mobile station even if the transmission quality is still adequate. This may happen when the global interference level would be significantly improved if the mobile station would be in contact with another cell. Computations and simulations show indeed that there is usually a "best cell" from the point of view of interference. This statement is especially true when power control is being used, since the cell corresponding to the minimum path loss will minimise the mobile station transmission power (ordered by the BTS), thereby statistically minimising the overall interference level. A handover triggered with the goal to optimise the interference level and not for the sake of the ongoing communication shall be referred to here as a "**confinement handover**". Such handovers result in a "confinement" in optimal geographical areas of the mobile stations which have a connection in a given cell, preventing them from wandering out of the optimal cellular coverage even if their connections are still of adequate quality. Confinement handovers may potentially conflict with local optimisation of the transmission quality, and should not be performed toward cells for which transmission quality is not correct.

A third kind of handover is referred to as the "**traffic handover**". It may happen that a cell is congested whereas neighbour cells are not. Such a situation happens typically when specific events lead to a very local geographical peak: fairs, sport events, and so on. Because the actual coverage of neighbour cells often overlap a lot, handing over some calls from one cell to a less congested one could temporarily improve congested situations. This kind of handover must be handled with great

care, since it is obviously in conflict with the confinement criteria. Traffic handovers will necessarily perturb cell planning and increase the level of interference in the surrounding area.

The *Specifications* handle the traffic handover in a different way from the other types of handover. The concepts of rescue handover and confinement handover, on the other hand, are not developed in the *Specifications*, but the following paragraphs will show how much this distinction can help operators and manufacturers when developing smart handover algorithms.

6.1.4.2. Handover Criteria

Depending on the purpose for handover, the criteria to be taken into account differ, but they always include some information to predict what will happen with and without handover, according to the destination cell.

The main criterion for rescue handover is the quality of transmission for the ongoing connection, both uplink and downlink. The best information would be an assessment of the transmission quality as perceived by both users. With digital transmission, the transmission error rate is a good quality indicator. The propagation path loss incurred by radio transmission is also of interest. Another piece of information, although of more marginal application, is the propagation delay. Transmission on the GSM radio interface cannot usually support a high propagation delay, and a connection can be cut if it becomes too big. The case may only arise in large rural cells. In GSM, all these measurements are available to the handover decision process. Both the mobile station and the BTS measure regularly the transmission quality and the reception level, from which the path loss can be inferred. The mobile station transmits its measurements to the BTS, at the rate of once to twice per second.

The key criteria for a confinement handover are the uplink and downlink transmission quality corresponding to each neighbour cell, were the mobile station to be in connection with that cell. Since this information is quite difficult to get (it may depend on the would-be allocated channel for instance), the handover process in GSM has to make do with only the path loss between the mobile station and a number of neighbouring cells. In reality, only downlink values are measured. by the mobile station, and the assumption is made that the path loss is equivalent in both directions.

The decision process for traffic handovers requires information on the load of each BTS, and this information is known by the MSCs and

BSCs. Traffic handovers differ quite a lot from the rescue and confinement handovers, because traffic reasons dictate the number of mobile stations to be handed over, in a given cell, but do not indicate which of these should be. The choice of the favoured (?) ones usually starts with those which are closer to be handed over for other reasons. Hence, the traffic handover relies on the other criteria and the corresponding measurements.

The algorithms for the handover decision and the choice of the target cell are not imposed by the *Specifications*. An example of such an algorithm is given as an annex to TS GSM 05.08, but operators and manufacturers have complete freedom to implement more (or less) sophisticated algorithms based on available parameters. In order to summarise, a list of the parameters to be taken into account in the handover decision process, is given hereafter:

- some static data, such as the maximum transmission power of:
  - ⇒ the mobile station
  - ⇒ the serving BTS
  - ⇒ the BTSs of neighbour cells;
- real time measurements performed by the mobile station:
  - ⇒ the downlink transmission quality (raw bit error rate)
  - ⇒ the downlink reception level on the current channel
  - ⇒ the downlink reception levels from neighbour cells;
- real time measurements performed by the BTS:
  - ⇒ the uplink transmission quality (raw bit error rate)
  - ⇒ the uplink reception level on the current channel
  - ⇒ the timing advance;
- traffic considerations, cell capacity and load, ...

Some of these parameters raise technical questions, for instance how does the mobile station reports its measurements, and how can a mobile station make measurements concerning neighbour cells with which it does not have a dedicated connection?

6.1.4.3. Measurements

Various aspects of the measurement process in GSM merit some explanations. One of them is how the measurements done by the mobile station and the base station must be transferred to a single point for treatment. This will be the first point we will look at. Another point is how the mobile station manages to make measurements concerning

neighbouring base stations. Finally, some of the difficulties of the specifications of what is measured will be addressed.

Measurement Reporting

In order to make efficient handovers, the rate at which measurements are refreshed should be as high as possible. In GSM, the minimum rate of reporting is once per second. The mobile station must report measurements, not only for the serving cell, but also for as many neighbour cells as possible which might be candidate target cells. This is true in particular for confinement handovers, since the mobile station does not have enough knowledge of the cell planning aspects to determine in all certainty which neighbour cell would be the best target cell. Such a choice depends not only on the path loss, but also on maximum transmission power, cell size, etc. In GSM, the mobile station can report on up to 6 neighbour cells in addition to the measurements relative to the serving cell.

The measurement reporting activity of the mobile station represents a rate of roughly 130 bit/s at least. This reporting is carried by messages on the small signalling channel associated with each TCH and called the SACCH, whose maximum capacity equals twice this rate. Hence, the refreshing rate may go up to two reports per second, if the SACCH is not used for other purposes in parallel. And here lies one reason why it was chosen to have a channel separated from the main channel. If the TACH (TACH = TCH + SACCH) multiplexing had not been done on a burst basis, but through a sharing of bursts (a few bits for the SACCH, the remaining bits for the TCH), then discontinuous transmission would have been useless because of the constant requirement for measurement reporting.

Neighbour Cells Measurements

The requirement for a mobile station to measure the reception level of neighbour cells, while carrying a call in a given cell, raises a number of technical issues.

A first issue is simply when can a mobile station perform these measurements. In GSM, it is possible for the mobile station to make these measurements while maintaining a connection, and without requesting the mobile stations to have two receivers, thanks to the TDMA scheme. The mobile station indeed measures the characteristics of the neighbour cells during the interval between the transmission of an uplink burst and the reception of a downlink burst. These intervals are of various lengths, depending on the dedicated channel type. The worst case corresponds to

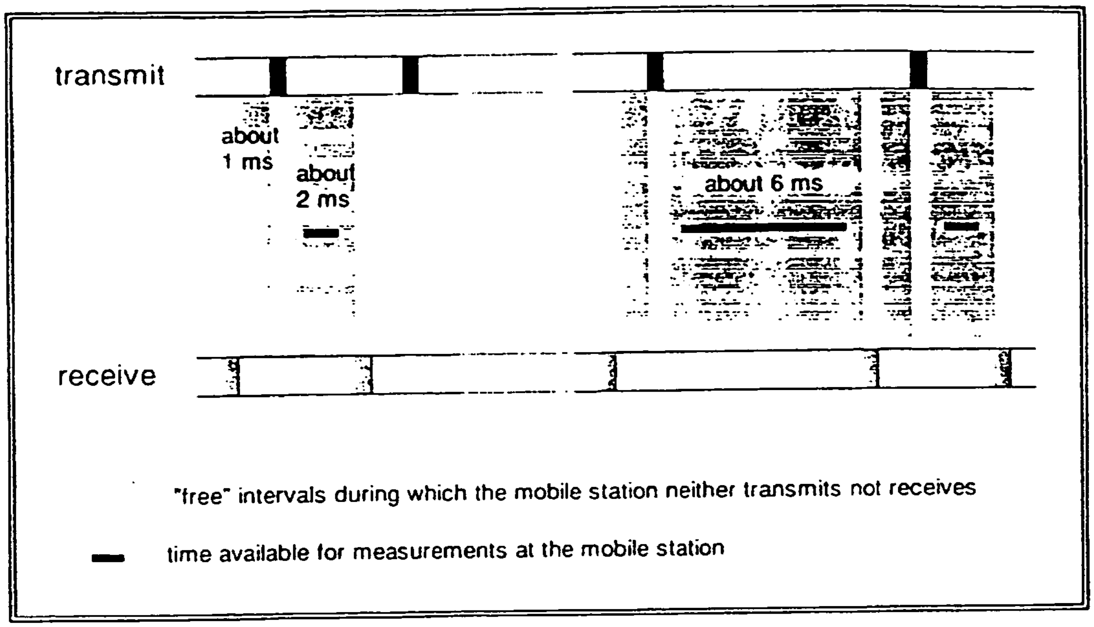


Figure 6.7 – Measurement intervals available at the mobile station

A TACH/F schedule leaves the mobile station with 26 very short intervals (2 BP – ε, # 1 ms), 24 "small" intervals (4 BP + ε, # 2 ms) and one "long" interval (12 BP + ε, # 6 ms) every 120 ms.

the TACH/F. The intervals, for each 120 ms period, are shown in figure 6.7 and are as follows:

- from end of reception to start of transmission:

  26 intervals of 2 BP –ε, where a BP (burst period) lasts for 577 μs and where ε represents the timing advance;

  these intervals are too short to be used for measurement.

- from end of transmission to start of reception:

  24 "small" intervals of 4 BP + ε and

  one "long" interval of 12 BP + ε. This interval exists thanks to the unused slot in the 26 slot cycle.

From these values, it is obvious that any efficient measurement scheme requires the mobile station to make measurements, not only during the long interval, but also during the small intervals. The measurement must be done on a frequency which is different from the one used for the preceding transmission burst or the next reception burst. This leads to a constraint on the frequency synthesis capability of the mobile station: either it is able to switch the reception/transmission frequency in less than 1 ms (thus leaving 300 μs for a measurement), or else it must have two frequency synthesisers. It can be seen that the

requirement for a multi-frequency synthesis capability of the mobile station does not derive from the frequency hopping alone.

Another technical difficulty emerges from the preceding discussion, concerning the infrastructure. As explained above, the schedule of the mobile station is very tight indeed; as a consequence, the mobile stations cannot be overburdened with the requirement of listening to a specific channel in neighbour cells, which in GSM would mean find one specific burst among eight in addition to switching to the right frequency. This has led to one of the most troublesome specifications (for BTS manufacturers!) of GSM. The chosen solution requires each BTS to emit continuously (i.e., in every burst period) on one frequency, moreover at a constant power level. This must be done regardless of the status of the corresponding slots, i.e., whether they are used by some active channel or idle, and necessarily without applying discontinuous transmission or power control. This continuity of the transmission allows the mobile stations in neighbour cells to make reliable measurements whenever they can, without any other constraint than their own scheduling.

The next problem which arises is how does the mobile station determine which frequencies it must measure? First, it is necessary for the mobile station to distinguish those frequencies which are used as beacon frequencies in other cells from other potentially interfering frequencies, in order to make sure which cell to measure. Second, it is more efficient for the mobile station to limit measurements to the beacon frequencies of neighbour cells. The latter point is solved by sending to the mobiles stations a list of frequencies to be measured. The first point is somewhat more tricky. It requires the mobile station to do a bit more than just measuring a received level.

The beacon frequency is the one carrying the synchronisation and frequency correction channels (SCH and FCCH) of the cell. A way for the mobile station to check that the channel it receives is actually a beacon channel (and not another channel of another cell transmitting on the same frequency) is to find if this frequency carries a FCCH. This checking helps in fact to meet another requirement as well: **the pre-synchronisation**. Because synchronisation is necessary before transmission, as explained in Chapter 4, a mobile station must get some synchronisation information from the cell it will be handed over to, at the latest during handover execution. Since handover optimisation was felt very important by those who drafted the *Specifications*, it was decided that mobile stations should acquire synchronisation with all cells on which they report measurements. Hence, a mobile station is constantly pre-synchronised with all cells to which it may potentially be handed over.

Pre-synchronisation requires that the mobile station decodes not only the FCCH, but also the SCH of the beacon channel. When can the mobile station perform these tasks? How can it be assured to find the FCCH given the short time it is able to spend for its measurements? This is where the "long" interval comes in. The reader will recall that, in the case of a TACH/F, there is one such "long" interval (lasting about 6 ms) every 120 ms, and more of them in the case of TACH/8s. As far as the TACH/F is concerned, the "idle slot" corresponding to this "long" interval, and which happens once in every 26 slot cycle, is indeed of critical importance for the pre-synchronisation process. This *idle slot* is indeed a misnomer, since the mobile station is anything but idle during this period. Looking for a FCCH requires indeed the possibility to listen to the potential beacon channel through a sufficiently long window; the FCCH cycle recurs with a period of 51 × 8 BP, whereas the *idle slot* recurs every 26 × 2 × 8 BP.

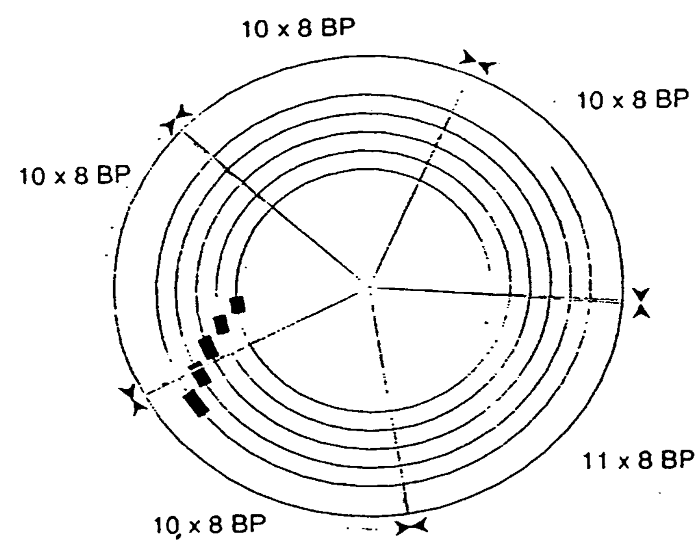


Figure 6.8 – Sliding measurement cycle

The interval between two successive FCCH bursts (including one of them) is at most 88 BP, and lasts most often 80 BP.
Since in 11 successive long intervals the mobile station will have been able to read more than 88 successive bursts,
it must have encountered an FCCH burst if the channel it monitors is a beacon channel.
The average time to encounter such a burst is roughly half of this value, i.e., 0.6 second.

This is where the arithmetic properties of the numbers 26 and 51 intervene: since these two numbers have no common divider, the two cycles are shifted in such a way as time passes that the *idle slot* has 100% chance of being aligned with the FCCH within 11 cycles. Figure 6.8 (previous page) shows that the FCCH cycle is shifted 8 BPs compared to the TACH/F cycle between two successive *idle slots*. Since a long interval lasts for about 12 BPs and therefore enables reception during about 9 BPs (the rest being used for frequency setting), 11 successive long intervals (i.e., about 1.2 seconds) are enough to ensure the mobile station with some opportunity to find an FCCH burst, as shown in the diagram. It is up to mobile stations to make good use of this opportunity!

In the case of the TACH/8, things are a bit different, because its recurring period is the same as the one of the FCCH. But the phasing between the two cycles can be anything. Figure 6.9 shows that there is

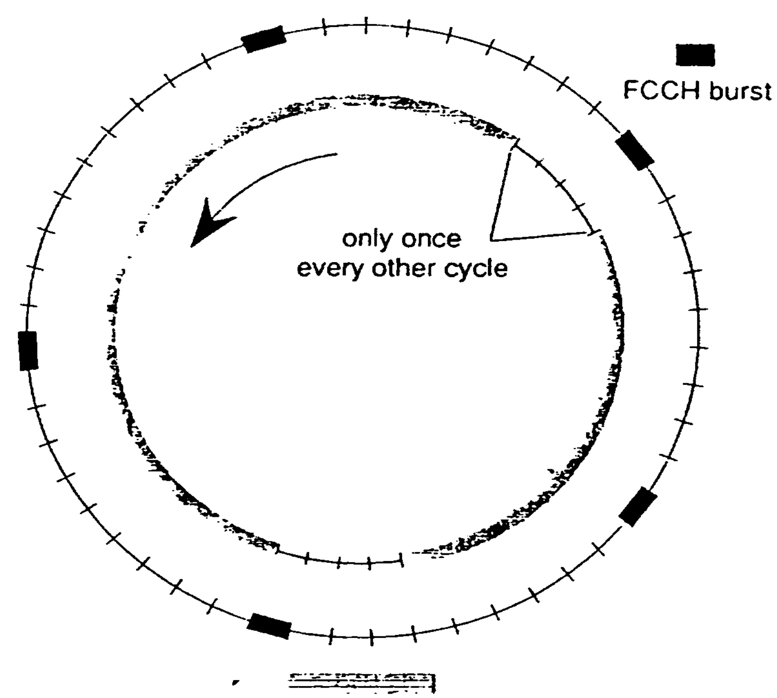


Figure 6.9 – TACH/8 cycle vs. FCCH cycle

Whatever the respective position of the TACH/8 and FCCH cycles, represented on the inner and outer circles, there is always an opportunity for an FCCH burst to fall into a measurement window of the TACH/8. The representation of the TACH/8s in this figure correspond to the scheduling of TACH/8s grouped by 8. The other case (TACH/8s grouped by 4) results in a slightly different but similar pattern.

necessarily a moment at which one of the 5 FCCH bursts in the 51 × 8 BP cycle can be listened to. The figure shows the case of TACH/8s not associated with the control channels, but the same reasoning applies for the TACH/8s associated with control channels.

Another question to be answered is which neighbour cell measurements does the mobile station report to the BTS among those cells measured, and how is this information packed into a small message? A measurement message can include up to 6 measurement results pertaining to neighbour cells. However, the mobile station may pre-synchronise with more than 6 neighbour cells. If this is the case, then only those measurements corresponding to the 6 cells it receives best are reported to the BTS.

Of course, this introduces a bias in the measurements collected by the network against the cells which are not among the first ones on the list. Their measurements are transmitted only when they are among the 6 best ones, and in consequence the ordering as established by the BTS by averaging several measurements may differ somewhat from the corresponding ordering by the mobile station. The measurement analysis process in the BSS ought to take into account the possibility of missing measurements; however, this potential error should not touch the best cells, which are of most interest.

The BSIC

It has been explained above how the mobile station is able to recognise that the frequency it monitors does correspond to a beacon channel. But there could be configurations where the mobile station is able to capture more than one beacon channel using a given frequency. This may happen when frequency planning must be done with very few frequencies, or at national boundaries. In order for the mobile station to be able to discriminate between the cells transmitting their beacon channels on the same frequency, a mechanism based on the BSIC (Base Station Identity Code, one of the major misnomers in GSM) has been introduced and warrants a detailed explanation.

The BSIC (Base Station Identity Code), which applies more properly to a cell, is *not* an unambiguous identification of a base station. Many cells bear the same BSIC, and moreover the normal practice is to allocate the same BSIC to neighbouring cells. So what is it? The BSIC is in fact a "colour code" (by reference to map colouring), allowing mobile stations to distinguish cells which transmit their beacon channel on the same frequency (see figure 6.10). For instance, when the radio spectrum available to a given operator is limited to, say, 2 MHz, frequency planning must cope with at most 10 frequencies. The best beacon frequency allocation scheme may not be able to avoid overlapping coverage in this case, and a mobile station will in some cases receive two beacon channels with the same frequency. A similar situation is also

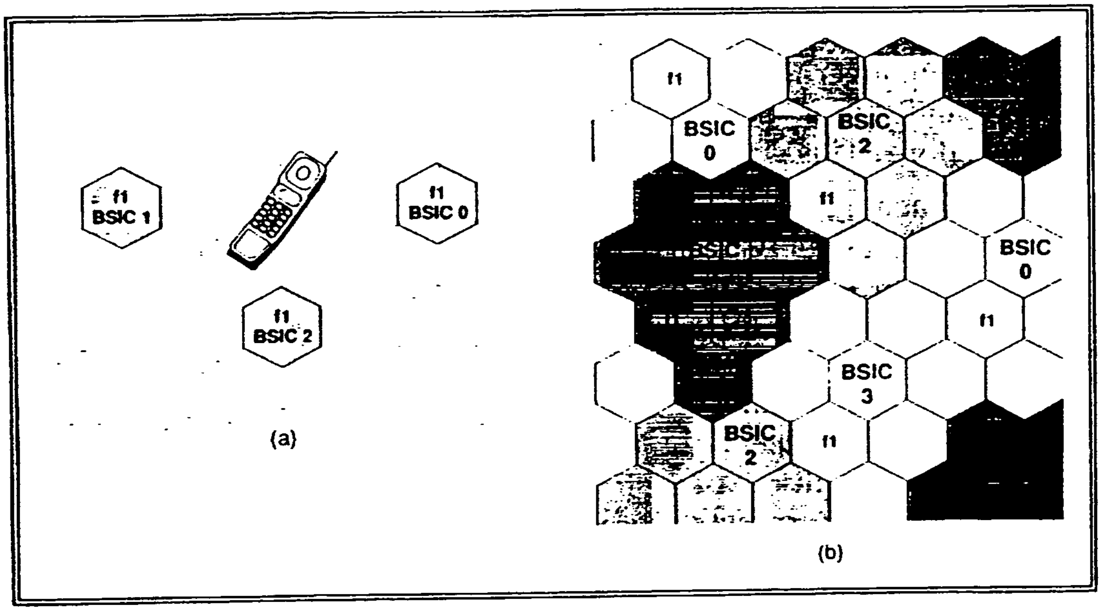


Figure 6.10 – Choice of the BSIC

The BSIC is a "colour code" allowing the mobile stations to distinguish cells which share the same beacon frequency (a).
When a regular reuse pattern is used for beacon frequencies, this pattern can be used as a typical basis for BSIC allocation (b).

frequent along national boundaries. Whereas inside a country the frequency allocation of different operators are disjoint, two public land mobile network (PLMN) operators on each side of the border will have some frequencies in common. Frequency usage co-ordination between operators helps, but cannot be enforced. In most cases a mobile station will still be in a position to receive the same beacon frequency transmitted by two base stations of different PLMNs. For all these reasons, a scheme allowing to distinguish cells using the same beacon frequency was deemed necessary. This is the role of the BSIC, a 6-bit code word broadcast on the SCH of every cell.

The BSIC intervenes in different cases, all related to the distinction between cells using the same beacon frequency:

- in order to perform neighbour cells measurements, the mobile station is provided with the list of frequencies to monitor. In the reporting message, the mobile station is required to indicate the BSIC for each beacon frequency on which it reports measurements. This implies that the mobile station has decoded the SCH of the beacon channels it measures, but this is not an additional constraint, since it must already be done for pre-synchronisation. The availability of the BSICs for measurement processing at the BSS enables it to check which cell has effectively been measured in the case of ambiguities.

- to prevent measurement reporting for cells toward which handover is precluded, a mechanism allows the network to indicate a subset of BSICs for which reporting should not be done. This screening mechanism is explained later.
- when a mobile station sends an access burst on the RACH of a cell, there is a risk that this random access be received by another cell than the one it was aimed at, if these cells use the same RACH frequencies and if their TDMA synchronisation do not fall too far apart. In order to avoid such spurious receptions, the RACH coded burst is "exclusive-ored" with the BSIC, so that only the right cell has a chance of decoding the burst successfully, based on the redundancy added to useful bits for checking correct decoding.
- when a mobile station in idle mode monitors neighbour cells, it regularly reads their BSICs; this provides the mobile station with a quick way to check whether the monitored cell is still the same or not. Such a check could also be achieved by decoding the broadcast messages containing the full cell identity, but at a greater cost.

The whole issue of BSIC planning consists in allocating different BSICs to cells using the same beacon frequency and between which overlapping coverage areas may exist. Inside a PLMN, BSIC planning is fairly easy; the matter is different at borders, when some co-ordination is necessary. In order to help this co-ordination, a tentative allocation of the first three bits of the BSIC has been introduced on a country basis, as shown in figure 6.11. This three-bit part has been named the "PLMN colour code" (or NCC, for Network Colour Code), a term which has been widely misunderstood. Its value is *not* a definite attribute of a PLMN, but a possibility for use near boundaries, and which can be overruled by bilateral or multi-lateral agreements between the concerned PLMNs. Indeed, nothing prevents all 64 values of the BSIC (including all 8 values of the "PLMN colour code") being used inside a PLMN when far away from any international boundary. Moreover, two PLMNs in the same country usually have disjoint frequency allocations, and can use without risk the same BSICs (and the same "PLMN colour codes"!).

The first 3-bit part of the BSIC is also the basis for screening measurement reporting. When two cells of different PLMNs using the same beacon frequency may potentially overlap near a border, it is normally best for a BSS to indicate to mobile stations that they should not report measurements concerning cells of the other PLMN, since these measurements would be worthless (since inter-PLMN handover is usually not performed), and might prevent the measurements of some cells in the

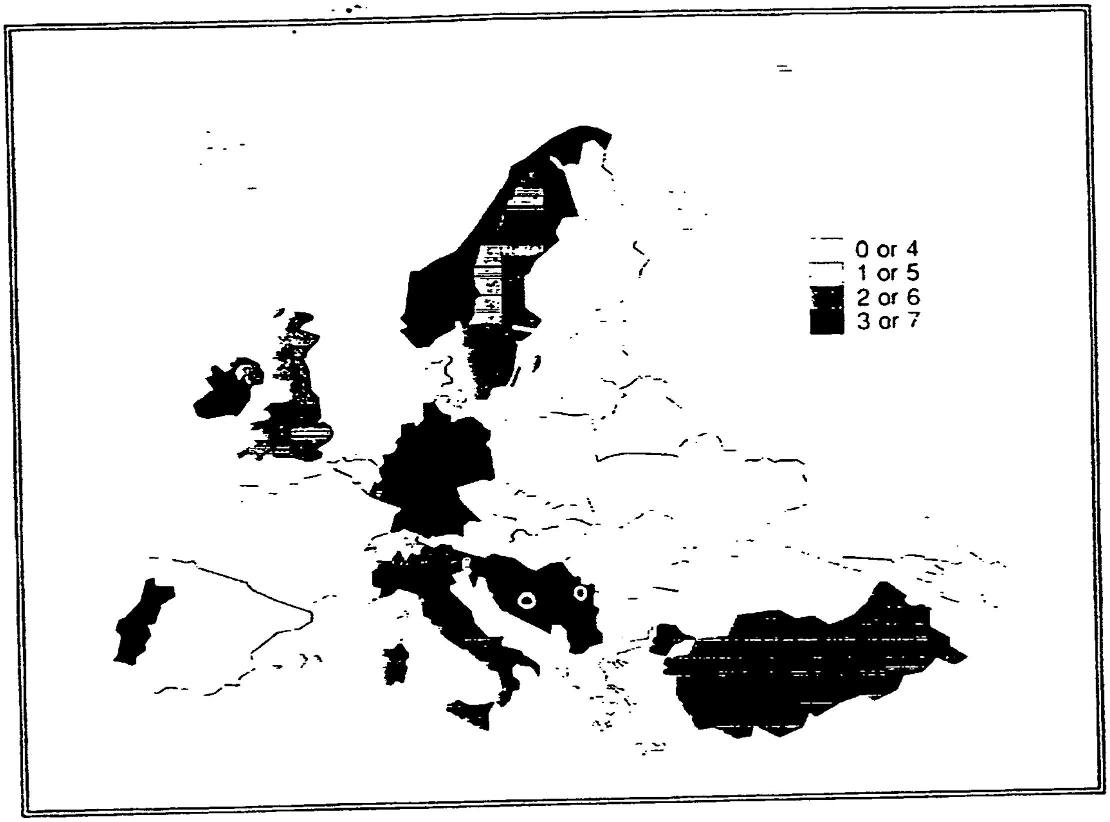


Figure 6.11 – A European PLMN code map

This tentative allocation of the "PLMN code" by country
is proposed to ease allocation of the BSIC near international boundaries;
but it does not represent a requirement.

right PLMN to be reported, since the number of reported measurements is limited to 6 neighbour cells. To that end, the information broadcast by each BTS includes an 8-bit screening indication, with one bit indicating for each of the 8 possible three-bit patterns whether cells with a BSIC starting with this pattern must be reported on. This screening indication is the *PLMN PERMITTED* indication.

The *PLMN PERMITTED* information happens to be broadcast on the BCCH. A current misunderstanding is that it influences cell selection (its name seems to carry such information indeed). This is not so: this information impacts only the measurement reporting and nothing else.

The Measurement Period

We will now look in some details to the measurement process itself. The various characteristics to be measured (reception level or reception quality) are continuous by nature, and affected by a significant level of "noise", that is to say statistical variations due to thermal or

industrial noise, Rayleigh fading, interference, and also due to the change of frequency introduced by frequency hopping. To be usable for decision-making, the measurements need to be filtered (e.g., averaged). The first step of the filtering is done in the mobile station. It is quite simple: it consists in taking the average value of each measured parameter for the duration of the reporting period, which will be described in the following paragraphs. The other point worth noting is that what is averaged is the raw bit error rate for quality measurements, and the logarithm of the reception level, or, in other terms, the reception level expressed in decibels.

A measurement report pertains to a given measurement period, that is to say the period during which the measurements were done. The duration of the measurement period is always equal to the periodicity of message transmission on the SACCH (i.e., 480 ms on the TACH/F, and around 471 ms for the TACH/8). On the TACH/F, the start of the measurement period starts a fixed time before the start of transmission. But on the TACH/8, things are different. The absolute position of the measurement period is the same for all TACH/8s of the same TN (timeslot number, indicating the position in the 8 BP cycle). Its relation to the time of transmission of SACCH messages therefore varies, depending on the TN. In addition, the position of the measurement period is different for TACH/8s combined with common channels and for those which are not. In all cases however, the uplink and downlink measurement periods are simultaneous.

A small subtlety in the definition of the measurement period comes from the quality measurement. With the interleaving scheme for speech, there is one speech block spread over two successive measurement periods in the case of a TACH/F. Quality measurements for this block belong to the second period. This problem should not have arisen for the TACH/8, where the aim was to design the measurement period carefully so that blocks are all entirely within a measurement period. However, there is one exception to this rule, since the third TACH/8 of TN 4 does not comply with this aim.

A simultaneity problem arises between the measurements done by the mobile station and those done by the BTS. The second ones are known by the BTS shortly after the end of the measurement period, whereas the measurement report from the mobile station for the same period is delayed by the message transmission delay, which amounts to roughly half a second on a TACH/F. In order to provide the BSC with reports which match, the BTS must then buffer its measurements until it has received the message from the mobile station. As seen by the BSC, measurements are received roughly one second after the start of the measurement period in the case of a TACH/F, and between a bit more

than half a second and close to one second for the TACH/8, depending on their position in the 102 BP cycle. The shorter delay in the case of a TACH/8 comes from the fact that each SACCH message is sent over 4 consecutive bursts, and not spread as for the TACH/F.

The Interaction with Discontinuous Transmission

The accuracy of measurements concerning the ongoing connection raises two problems, one in relation to discontinuous transmission, the other linked with power control and frequency hopping when using the beacon frequency.

When discontinuous transmission is applied, some slots belonging to a channel may not be used for effective transmission. This is indeed the goal of discontinuous transmission. But then measurements on these slots will obviously report a low reception level, and a bad quality. Even more annoying, one could imagine a measurement period, or a succession of periods, with no transmission at all, hence some difficulties for the processes which rely on these measurements. To circumvent these problems, the *Specifications* impose that at least 12 bursts are sent within each reporting period. These bursts amount to the systematic use of the SACCH for effective transmission (4 bursts constituting a coding block), and 8 bursts on the TCH itself. On a TCH/8, which corresponds to exactly 8 bursts per measurement period, this leads to the consequence that discontinuous transmission is not applied. On a TCH/F, this means that at least one block per measurement period must be sent (a block being interleaved over 8 half-bursts). For speech, this block contains a silence descriptor frame (SID frame) refreshing the comfort noise characteristics.

In addition to this minimum transmission rule, the *Specifications* require the BTS and the mobile station to report *two* sets of measurements concerning the connection: "full" measurements, done on all slots which may be used for transmission in the reporting period, and "sub" measurements, done only on the mandatorily sent bursts and blocks. On a TACH/F, the second set is less accurate than the first one when discontinuous transmission is not used, because averaging is done on a smaller set (for instance reception level is averaged on 12 bursts instead of 100 bursts). On a TACH/8, the two measurements are evidently identical, but are nevertheless both sent for uniformity. Finally, both the BTS and mobile station report for each measurement period whether discontinuous transmission was effectively used or not (or in another terms, whether all bursts were effectively transmitted), thus enabling the processes using the measurement to discard the "full" measurements provided by the other end when applicable.

The PWRC Indication

Another problem with measurements (and the last point on the topic) is in fact a consequence of a combination of different independent details of the *Specifications*. First, it is allowed that a frequency hopping TACH uses the beacon frequency as one of its frequencies (of course, not on TN 0). Second, power control can be applied on the downlink TACH. Third, the beacon frequency must be transmitted with a constant transmission power, because of the measurements performed by mobile stations of neighbour cells. The result for the channels under consideration is that power control applies only to a subset of the bursts, whereas other bursts (those using the beacon frequency) are sent with a fixed transmission power. This leads to inaccurate reception level measurements. In order to alleviate this problem, the mobile station is requested in such cases not to take into account the slots falling on the beacon frequency in the reception level estimation. This is controlled by an indicator, the *PWRC* indicator (originally called power control indicator, a misnomer), sent on a connection basis to the mobile station. This indicator should be set if the following conditions are all met: the channel hops on at least two different frequencies, one of those frequencies is the beacon frequency, and downlink transmission power control is in use.

6.1.5. POWER CONTROL AND TIMING ADVANCE

Most of the functions needed for transmission over the radio interface correspond to transformations of the signals representing the data to transmit. They were the object of Chapter 4. Two functions, the management of the transmission power and of the timing advance, lie somewhere between these pure transmission functions, and transmission management functions. As the transmission functions, they are performed continuously and relate heavily to physical features, but they share the use of signalling means with the transmission management functions. They are studied here, because of this usage of signalling means, and also because they are, as we will see, deeply entangled with the procedures managing RR-sessions. This comes from the initialisation of these processes, which must take place any time the channel is changed.

6.1.5.1. Power Control

Power control refers to the possibility to modify within some range the transmission power on the radio, both (but independently) for the mobile station and the base station. Power control shares a common goal with discontinuous transmission: improving spectral efficiency and,

although to a lesser extent, battery life for the mobile station. When one side is received too well by the other, it becomes advantageous to reduce its transmission power by such an amount as to keep a similar quality level on the communication, while decreasing the interference caused on other calls in surrounding areas. In a system such as GSM, the full gain in spectral efficiency can be obtained with a small power range, and the 20 dB minimum case specified by the *Specifications* is more than sufficient for this purpose.

In GSM, both uplink and downlink power control may be applied independently one from the other; furthermore they are applied independently with each mobile station. The range specified by the *Specifications* for uplink power control lies between 20 and 30 dB, by steps of 2 dB, depending on the mobile station power class. An example is given in figure 6.12. The range used for downlink power control is manufacturer dependent and may be up to 30 dB, also by steps of 2 dB. The control of the transmission power is a network option, i.e., the operator may choose to apply it or not, in one direction or in both. All mobile stations, though, must support the feature, thereby allowing power control to be really efficient when utilised.

Power control on both directions is managed by the BSS. The transmission power of the mobile station is chosen by the BSS, and

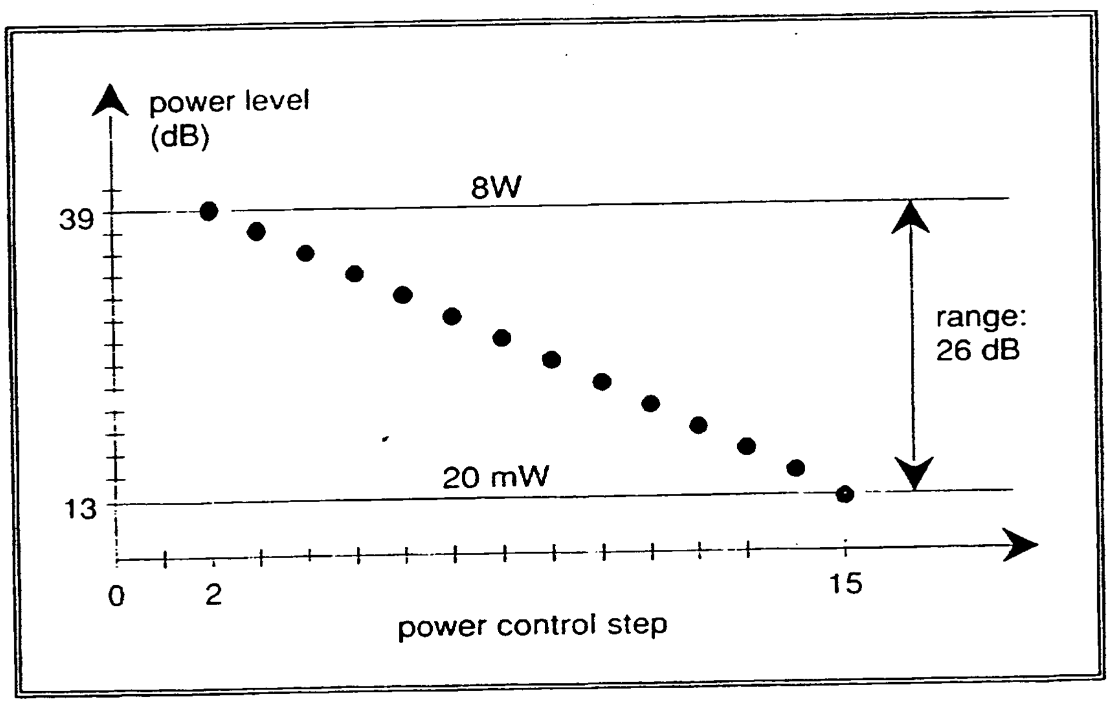


Figure 6.12 – Power control steps for a class 2 GSM900 mobile station

Power can be controlled by steps of 2 dB on a range going from 0,2 W (13 dBm) to the maximum MS power (here 8W = 39 dBm).

commands to regulate it are issued to the mobile station. The BSS computes the required MS transmission power through reception level measurements performed by the BTS, taking into account the MS maximum transmission power as well as quality measurements done by the BTS; this last parameter helps to ensure that transmission quality is kept above some acceptance threshold. For the downlink direction, the BTS transmission power is also computed by the BSS for each connection, based on the measurements performed by the mobile station and reported regularly to the BTS.

Inside the BSS, the split between BTS and BSC is basically an option for the manufacturer. The specification of the Abis interface as found in the *Specifications* is basically adapted to the implementation of power control in the BSC, but implementation in the BTS is possible. The detailed procedural means for the latter case are not specified in the *Specifications*, but a number of "place-holders" in the BTS-BSC protocol allow manufacturers to specify and implement part or all of the power control management in the BTS.

At the start of a connection, the initial value of the transmission power (both for mobile station and for BTS) is chosen by the BSC. In the case of an initial assignment, the information available to choose this power is at best very small: it consists in the reception level of a single access burst, which is necessarily of limited accuracy. Therefore, in

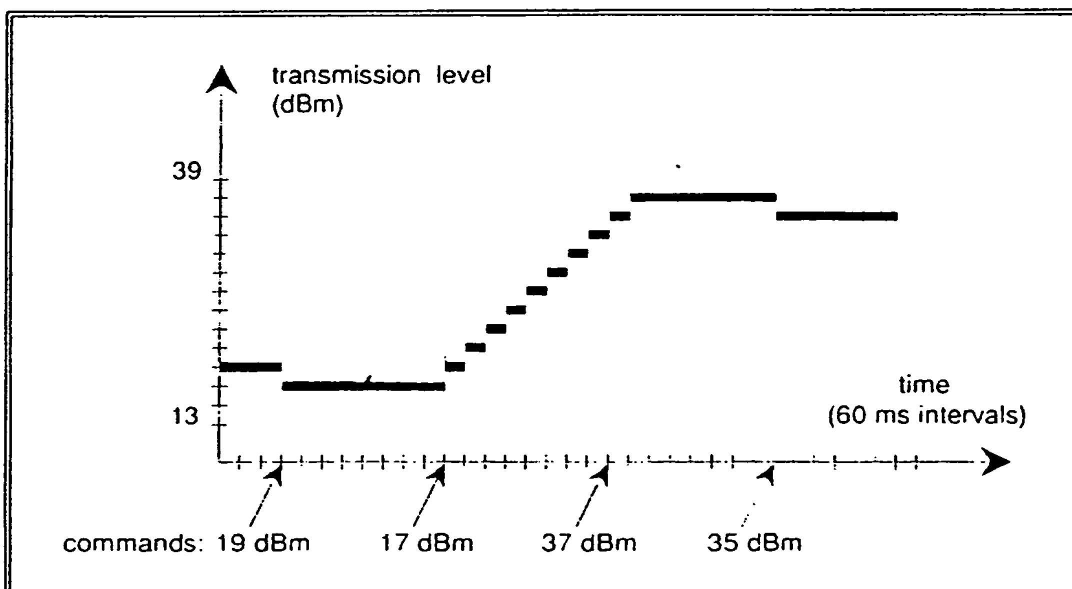


Figure 6.13 – Transmission power adaptation

The transmission power is adjusted by steps of 2 dB,
recurring not more often than every 60 ms.
A high jump in the power control commands will therefore be answered gradually.

GSM, the initial power level to be used by a mobile station for the first messages sent on the new dedicated channel is fixed on a cell-per-cell basis, and is the same level as used for sending random access bursts. The value of this level is broadcast on the BCCH, to be known by all mobile stations before any access attempt. A mobile station whose maximum power level is below the broadcast value shall simply use its maximum power level instead. For subsequent channel connections, the MS transmission power to be used when accessing the new channel is specified by the BSC, either using a default cell value (typically the case for an incoming handover between different BSCs) or based on the knowledge concerning the previous connection (this capability could be used, e.g., at subsequent assignment).

Except at the start of a channel connection, a command to change the transmission power does not trigger an immediate transition to the ordered value in the mobile station. The maximum variation speed is of 2 dB each 60 ms (see Figure 6.13). This means in particular that a high jump (more than 12 dB) will not be terminated when the next command arrives.

The basic procedural requirements for power control, as shown in figure 6.14, include the following:

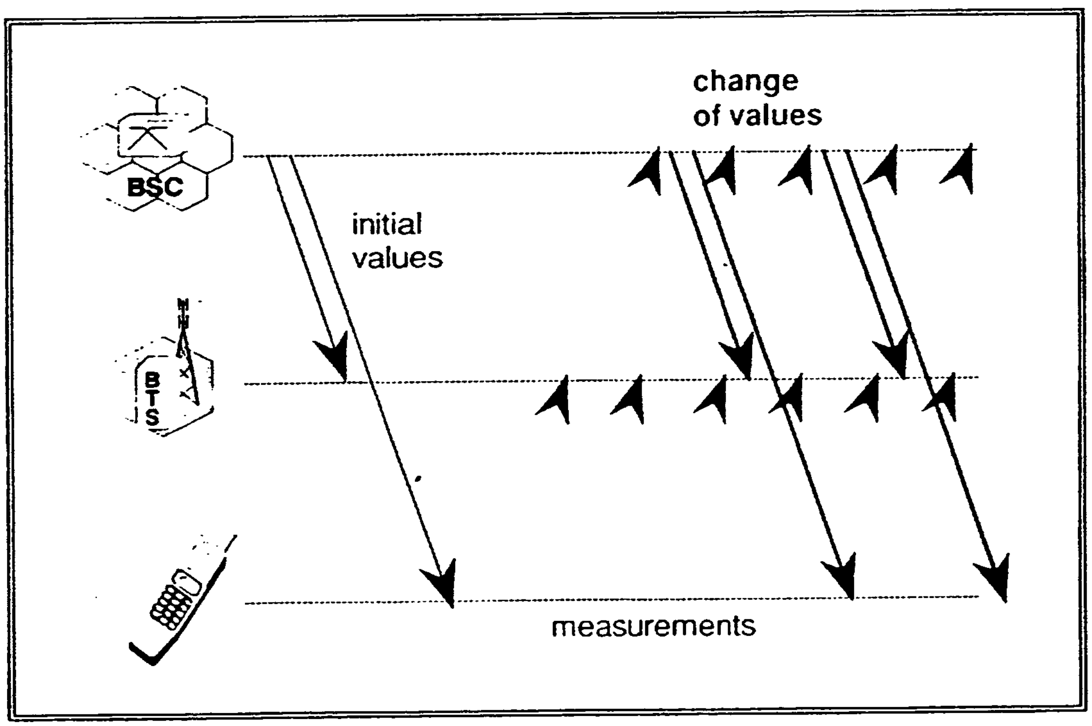


Figure 6.14 – Procedural needs for power control

Mechanisms to report radio measurements, to set-up the initial power value, and to control power by steps are included on the Abis and the radio interfaces.

- some means for the mobile station to transmit measurements (the same SACCH procedures are used for handover preparation) up to the BSC, even though some measurements may stop their journey at the BTS to be "pre-processed";
- some means for the BSC to command MS and possibly BTS transmission power;
- some means for the BSC to indicate to the mobile station the initial power level value to be used at initial assignment, as well as at each subsequent channel transition;
- some means for the BSC to indicate to the BTS the initial power level value when a channel connection is initialised.

6.1.5.2. Timing Advance

The time division multiplexing scheme used on the radio path of GSM is such that the BTS must receive signals coming from different mobile stations very close to each other. In order to reach this goal despite the propagation delay incurred by the return trip from BTS to mobile station, and taking into account that guard times between bursts have been chosen very small for spectral efficiency, a mechanism to compensate for the propagation delay is necessary. To this avail, the mobile station advances its transmission time relative to its basic schedule, which is derived from the reception of bursts, by a time indicated by the infrastructure, the timing advance.

Once a dedicated connection has been established, the BTS continuously measures the time offset between its own burst schedule and the reception schedule of mobile station bursts. Based on these measurements, it is able to provide the mobile station with the required timing advance and does that on the SACCH at the rate of twice per second.

The timing advance can take values from 0 to 233 μs, which is enough to cope with cells having a radius of up to 35 km without any other special scheme and given the speed of light. This limit comes from coding considerations (the timing advance is coded between 0 and 63, with the bit period as the unit, hence 233 μs), but there are more important hidden limitations. A first point is the guard time for access bursts, which in practice limits to about 220 μs the possibility for the

initial propagation time measurement. The other point is that some minimum time is needed between the end of the reception of a downlink burst and the beginning of the transmission of the next uplink burst, in order to allow the implementation of mobile stations with the same frequency synthesiser for emission and reception.

Even in rural or low-density areas, good coverage quality will in practice require cell radii smaller than 35 km. However, there are cases when larger cells would be useful. This holds in particular for the coverage of inshore coastal areas, where high antennas (e.g., on lighthouses) could be in sight of boats more than 35 km away on sea. Such uses are indeed possible, at the expense of the number of channels per MHz. The trick consists in obtaining a huge guard time (more than 580 µs) by using only every second burst. In such cases, only the channels of even TNs can be used (since TN 0 must be used for the BCCH, odd TNs will not do). This feature requires a specific reception processing in the BTSs.

Upon establishment of a new dedicated connection, the timing advance control process must be initialised. This happens at each initial assignment, subsequent assignment or handover. Depending on the case, the mobile station and the infrastructure do not always have the same amount of information to assess the new timing advance, and the initialisation method varies accordingly. The different cases will now be examined one by one.

a) Both Mobile Station and Network can assess the Timing Advance Beforehand

This situation happens upon subsequent assignment. The mobile station simply uses on the new channel the old value of the timing advance which was ordered by the BTS on the previous channel. On the infrastructure side, the BTS device in charge of the old channel is aware of the last ordered timing advance, but there is no communication means between BTSs. The BTS is indeed not aware that the old and the new channel connection concern the same mobile station. The BSC is kept informed of the last ordered timing advance value, and it is therefore able to transmit it to the new BTS device when activating it. The timing advance control process just resumes on the new channel with the same values as on the previous channel.

There exists another case where both mobile station and network are able to assess the timing advance beforehand. This happens at

handover between synchronised cells which are collocated. However, the *Specifications* treat this case as if the cells were synchronised, but not collocated (see next paragraph). A more efficient scheme could have been to give an indication to the mobile station that the old and new cells are collocated, and such an improvement may indeed be made in the future.

b) Only the Mobile Station can assess the New Timing Advance Beforehand

This case arises when handover is performed between two synchronised cells which are not necessarily at the same location. The mobile station is then able to measure the difference between arrival times of bursts coming from the two BTSs. It must indeed do so for pre-synchronisation requirements (see page 333). This arrival time offset is a combination of the transmission time offset between the two BTSs (which does not depend on the location of the mobile station), and of the two propagation times, as shown in figure 6.15. Therefore, if the mobile station is given the transmission time offset between the two BTSs (which is zero, by definition, for synchronised cells), it is able to derive

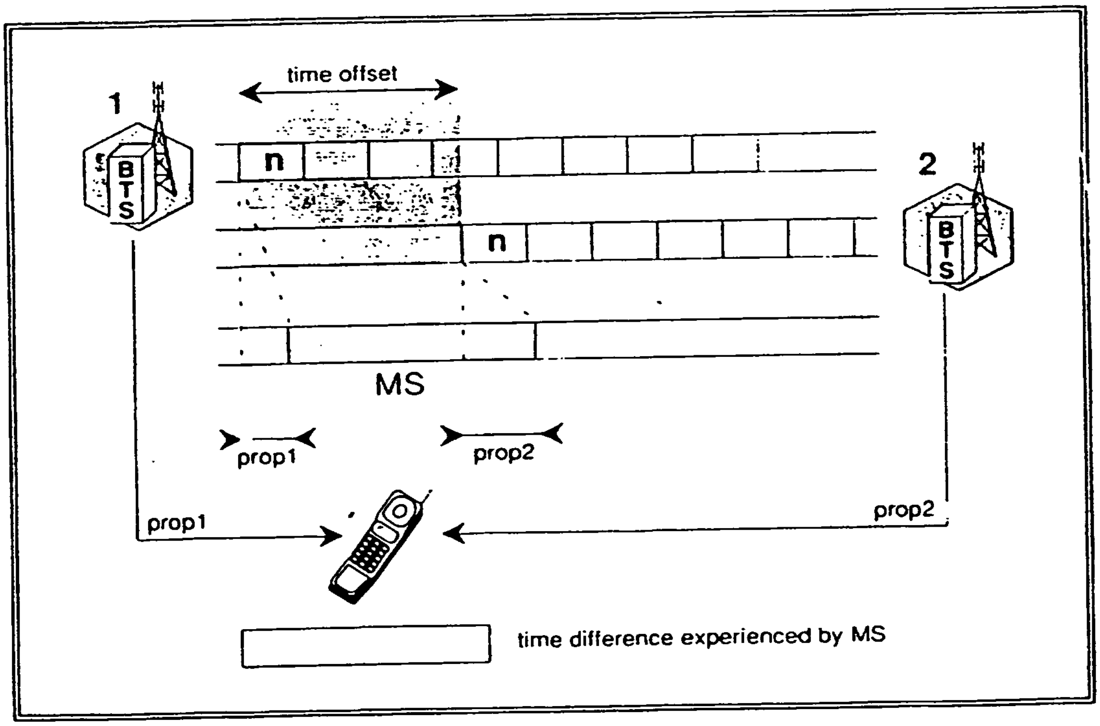


Figure 6.15 – Time offset between two BTSs

As seen by the mobile station, the arrival time difference between bursts coming from two BTSs is made up by the difference of the propagation times, plus the offset between clocks of the two BTSs.

the difference in propagation times, and therefore to calculate the new timing advance to be applied.

The indication that two cells are synchronised is therefore enough for the mobile station to assess the new timing advance, as follows (based on the notations of figure 6.15):

$$\boxed{TA2 = TA1 - 2\,(prop1 - prop2)}$$

On the infrastructure side, the new timing advance cannot be computed, except in the case of collocated and synchronised cells. A possibility would have been to wait for the mobile station to indicate this value (we will see later that the mobile station indicates back the value of the timing advance it uses, at least once per second). Nevertheless, to allow for a possible slightly faster initialisation on the new channel, the handover procedure between synchronised cells includes the means for the new BTS to assess the propagation time with the mobile station. To this avail, the mobile station starts transmission on the new channel with a few access bursts sent with a null timing advance, before switching to normal transmission.

c) Neither Mobile Station nor Network are able to assess the New Timing Advance Beforehand

At initial assignment, or at handover between two cells which are not synchronised, no information can be used by either side to predict the timing advance. Signalling messages are different, but the timing advance initialisation process is very similar in both cases. The mobile station is forbidden to transmit normal bursts until it knows the new timing advance to apply. Because the BTS (which is the entity which decides on the timing advance) must receive something from the mobile station in order to assess the propagation time, the mobile station is required to send access bursts to the BTS with a null timing advance. When the BTS receives such a burst, the reception instant is a measure of the double propagation time and the BTS can derive the value of the timing advance, which it sends to the mobile station in a signalling message. From the moment it receives this message, the mobile station is able to start correctly transmitting normal bursts. This exchange lengthens the duration of the handover procedure between asynchronous cells compared to the synchronised case described earlier. Moreover, the "asynchronous" handover leads to a longer communication interruption than its "synchronous" parent.

6.1.6. Radio Channel Management

So far, we have seen how individual RR-sessions are managed. Though RR-sessions are independent, they share the same pool of resources, in particular the radio channels. In this section we will look at the management of the radio channels in a cell as a whole, dealing with such problems as the configuration of the channels and the channel allocation strategy.

The management of the set of channels to be used in each cell includes two main aspects. On one side, the set of channels of each cell must be determined and the machines need to be configured accordingly. This "long-term" aspect is the cell channel configuration management. On the other side, the channels go through allocation/release cycles following the communication needs of the mobile stations. This "short-term" aspect is the dedicated channel allocation management. Both channel configuration management and channel allocation are the responsibility of the BSC. The MSC only intervenes to indicate which type of channel a given communication requires, whereas the BTS executes different related tasks, but always under control of the BSC. Both areas have strong impact on the procedures used over the radio interface and the Abis interface and will be described here.

6.1.6.1. Cell Channel Configuration

The channel configuration of a cell is the list of channels defined at a given time to be used in the cell. A typical cell configuration includes a set of common channels to support mobile stations in idle mode and initial mobile station access (a BCCH, a PAGCH and a RACH) and a set of traffic channels (TACHs of various rates, including what the *Specifications* call TCH/F, TCH/H and SDCCH) for carrying signalling and user data. The channel configuration of a cell may change in time. These changes may have various degrees of impact on traffic management, i.e., on the allocation and release of channels used for specific communications.

On one side, some modifications are related to the evolution of the whole network, for instance a capacity extension to cope with an increasing traffic density. Such changes are clearly within the scope of network operation. However, because network operators appreciate the possibility to handle such changes without disturbing the existing ongoing communications, mechanisms have been introduced in the traffic

management area for this purpose; these mechanisms will be described here.

Configuration of the Access Channels

Depending on the full spectrum capacity of the cell, usually estimated in terms of number of frequency slots, the capacity requirement for access channels (RACH, PAGCH) will vary. The *Specifications* cater for five different access channel capacities for a given cell, which are summarised in table 6.2. These access channel capacities correspond in terms of radio consumption to a range from the equivalent of a half-rate traffic channel to the equivalent of 4 full rate traffic channels.

Since the mobile stations are assumed to be able to listen to only one unit of spectrum usage (the equivalent of a TACH/F, i.e., one slot every 8 time slots) at a given moment, they must be distributed among 1 to 4 population groups, depending on the capacity of the access channel structure. Mobile stations find the information about the applicable structure in the messages broadcast on the BCCH. These messages are sent on every carrier unit used for common control channels (of TN 0, 2, 4 and 6 as applicable).

The access channel configuration may change in time, and detailed procedures are provided in the *Specifications* to cope with such dynamic changes and ensure that the transition period between two stable configurations is as limited as possible. This impacts mainly the listening to the PAGCH in idle mode.

CCCH capacity (equiv. in TACH/F)	**number of MS groups**	**RACH burst rate** (bursts per second)	**PAGCH message rate** (messages per second)
1/2 (*)	1	114.7	12.7
1	1	216.7	38.2
2	2	433.4	76.5
3	3	650	114.7
4	4	866.7	152.9

Table 6.2 – Access channel capacities

Depending on the capacity to be offered by a given cell, the access channels can be configured in several different sizes.
(*) the other half can be used only for 4 TACH/8.

The corresponding procedural aspects will be covered in the section dealing with paging procedures.

Organisation of the PAGCH

On each CCCH "unit" to which the mobile stations are able to listen, the downlink paging and access grant channel is organised in two parts:

- several "paging sub-channels", in a one-to-one relationship with sub-populations of mobile stations, on which initial assignment messages can also be sent.
- possibly a sub-channel reserved exclusively for assignment messages;

This PAGCH configuration is indicated to the mobile stations in messages broadcast on the BCCH, in order for mobile stations to determine where to listen for their own paging messages. The PAGCH configuration may change dynamically, and a mechanism is defined in the *Specifications* to enable such a change while avoiding the risk for mobile stations to lose paging messages during such changes. Both the corresponding contents of the broadcast messages and the allocation of paging messages to sub-channels are controlled by the BSC.

These procedural aspects will be described respectively in the sections dealing with general information broadcasting and with paging procedures.

Traffic Channels Configuration

Another point in the area of cell channel configuration management is the possibility to modify dynamically the set of traffic channels to meet the demand more closely. For example, the resource used at a given moment in time for a TACH/F can also be used for 8 TACH/8 at some other time. This kind of choice can be under control of the operation and maintenance sub-system (as a result of medium or long-term traffic analyses), or alternatively may be fully implemented in the BSC, so that the allocation process of the BSC would perform the conversion when needed. For instance, if a TACH/8 is needed when none is free but if a TACH/F is available, the latter could be converted into a pool of 8 TACH/8s instead of rejecting the request. The *Specifications* leave complete freedom to the operator/manufacturer to choose the implementation anywhere between these two extremes.

These functions do not impact directly the communications in progress, and are totally internal to the BSC. Hence they are not visible on the connection management protocols.

Changes in the Frequency Configuration

The previous paragraphs have dealt with changes in the functional configuration of channels, within a given pool of time/frequency resources. But that is not the end of the story; the frequency slots allocated to a cell may change dynamically in time, even though this is presumably not a very frequent event.

In the case when only single-frequency channels are used, a change in the frequency allocation of the cell will impact the communications making use of any suppressed frequencies, but each such communication can be handled independently from the others. However, when frequency hopping is employed, a frequency is used in a very tightly co-ordinated way by several connections in the cell at the same time. Any change in frequency affecting a given connection must be precisely coupled with similar changes to other connections in order to keep the non-interfering properties of the channels. These changes must happen in a synchronous way. With this objective, specific mechanisms have been introduced in the *Specifications* to enable the synchronised modification of the frequency allocation of many channels.

These mechanisms include the possibility to order a precisely timed change of the frequency parameters to the mobile stations and to the BTS for all impacted connections, as well as to have precisely timed channel assignments (whether initial assignments, subsequent assignments or handovers). The indication of the instant of change relies on the cyclic numbering scheme of TDMA slots, which has a period of about 3 and a half hours and allows an accuracy of microseconds.

As seen from the mobile station, these changes appear simply as changes of channels to be performed at a defined instant. On the BSS side, the matter is somewhat more complex. The aim is to synchronise the behaviour of several mobile stations and of the BTS, using signalling means which are by nature asynchronous and subject to losses due to transmission errors. The operation must be performed in several steps.

First, the general decision to perform a frequency change comes from the operation and maintenance sub-system, for such reasons as the setting up of new hardware, or the need for removing some for maintenance, or due to observations of unplanned interference. The

decision to modify the frequency organisation of a cell is notified to the BSC, which is then in charge of co-ordinating mobile stations and BTSs to reach the new coherent configuration.

The first step for the BSC is then to determine the transition instant. This instant must not be too far away, to avoid introducing ambiguities from the cyclic numbering scheme. However, it must not be too close, in order to ensure that all concerned mobile stations have either received the command or have had their ongoing connection released. The concerned mobile stations are those engaged in transmission on a channel whose frequency parameters are affected by the change, i.e., those to which such a channel has been allocated before the actual transition time: a timed transition order must be sent to them.

There remains the case of the new allocation of channels before the actual transition time: such procedures are started, but with an indication that the mobile station will go on the new channel only at the transition time. Thus, all mobile stations involved in a connection on one of the impacted channels, as well as the BTS, perform the transition when the transition time occurs, and a normal situation is restored.

To summarise, the change of frequency configuration impacts all the assignment procedures, and require a specific procedure to deal with existing connections, called the *frequency redefinition* procedure. The corresponding details will be found in the sections dealing with each of these procedures.

6.1.6.2. Dedicated Channel Allocation

The second component of the management of the radio channels seen as a set is how the dedicated channels (TACH/8 and TACH/F) are chosen when allocated to an RR-session. As seen by the infrastructure, dedicated channels are at a given moment either allocated to the use of a mobile station, or part of a pool of idle channels from which a channel is drawn when a new need appears. To summarise the previous sections, such a new need may appear in three different conditions:

- at initial channel assignment, when a mobile station in idle mode has some communication needs, for instance because the user wants to set up a call, or because location updating must be performed;
- at subsequent assignment, when what the communication needs does not correspond any more to the type of channel it is allocated, for instance when a TACH/8 was allocated at initial assignment and a speech call needs to be connected;

- at handover, when the movements of the user or the variations of the interfering level result in a situation where the connection would be better on another channel, often through another cell.

Allocation Strategies

From the mobile station point of view, these various kinds of channel assignments are simply orders to start transmission and reception on specified channels. From the infrastructure point of view, the allocation of a dedicated radio channel involves two steps, first the choice of the channel to use, second the actual transition. The choice of the allocated channel lies entirely within the responsibility of the BSC. Sophisticated algorithms can be designed to try maximising the total amount of traffic which can be served with a given amount of resources, while maintaining a reasonable fairness level in the granting of requests.

Allocation optimisation includes several aspects. A first one is related to the relevance of the type of channel which is allocated for the effective need. This leads to a real problem in the case of initial assignment, since very little information is available at the BSC to choose the type of channel (the mobile station gives only a rough description for its reason to access in the initial access request message). A typical example concerns the setting up of a call: a TACH/F will be required to transmit user data, but a TACH/8 using 8 times less resources would be enough until the correspondents have begun conversing. Several strategies can be chosen, which can be grouped under the three following categories, as also shown in figure 6.16:

- **Very Early Assignment** consists in allocating a TACH/F at initial assignment, when it is probable that the requested connection will need such a channel;
- **Early Assignment** consists in allocating a TACH/8 initially, then subsequently allocating a TACH/F as soon as it is known for sure that this type of channel will be required;
- **Off Air Call Set Up** (OACSU) consists in allocating a TACH/8 initially, then waiting until the called party has answered before attempting the subsequent assignment of a TACH/F.

These different methods each have their pros and cons, which fomented many debates over the specification years of GSM. OACSU differs from the two other methods in that it provides the users with a different grade of service. The correct channel may be allocated a noticeable amount of time after the called party has answered, depending on the availability of channels at this instant. OACSU therefore requires

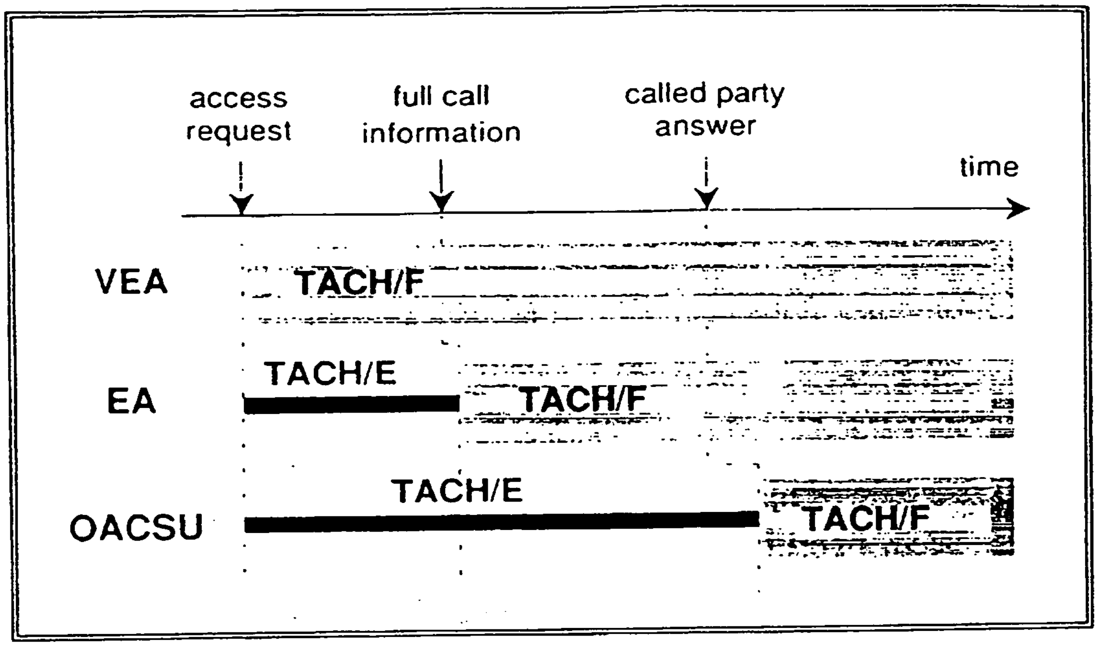


Figure 6.16 – Assignment strategies at call set-up

Very Early Assignment (VEA), Early Assignment (EA) and Off-Air Call Set Up (OACSU) represent three different allocation strategies, between which operators may choose.

some announcement to the called party, who may otherwise wonder why the phone rang at all! The grade of service which results is considered by many operators as unacceptable, though opinions diverge on the issue. On the other hand, OACSU is certainly the most efficient of the three schemes in terms of resource usage. The dilemma lies (as often) between efficiency and user comfort.

The main drawback of early assignment compared to very early assignment consists in an increased call set-up time, with no real gain in terms of channel usage in the case when a TACH/F is actually needed (it should be remembered that a measure of channel usage must take into account, not only the channel size (spectrum consumption), but also the usage duration, and that signalling exchanges are quicker on a TACH/F than on a TACH/8). On the other hand, very early assignment is very inefficient if a TACH/F is allocated when not necessary, e.g., for location updating. In this case, the amount of time during which the connection must be kept is mainly determined by the duration of signalling exchanges between infrastructure entities, and one cannot expect a significant reduction by using a larger channel. Therefore, very early assignment is of interest only if enough information on the use of the required connection is available to the network before initial assignment (which is not the case in the phase 1 version of GSM).

Whatever the allocation strategy, there are cases where no adequate resource is available when needed. The network may then apply one of two strategies: either the request is rejected, relying on its originator to possibly retry later, or the request is put aside to be served when a suitable channel becomes free. The latter strategy is referred to as "queuing", although there is not necessarily a queue in the strictest sense of the word.

Queuing

The interest of queuing varies with the conditions in which it is applied. At initial assignment, the repetition scheme put in place to cope with losses due to collisions or bad propagation conditions reduces the interest of queuing to nothing. Using queuing at initial assignment could even have adverse effects, since a request which is not answered in a short time will be repeated by the mobile station (the normal reason for no answer being transmission loss), and this could lead to very inefficient multiple assignments.

Queuing could of course have been made possible at initial assignment, by introducing a "please wait" acknowledgement to the mobile station request, but this has not been introduced for simplicity reasons. The *WAIT INDICATION* element appearing in a message answering a request for access should not be mistaken for such a scheme; it is in fact a temporary rejection preventing the mobile station from making new attempts for some time.

The major interest of queuing is indeed found in the case of subsequent assignments. The only resulting drawback is the lengthening of the call set-up time (this is perceived by the calling party when early assignment is used, and by both parties when OACSU is used). However, this is to be balanced with a rejection of the call, probably more ill-perceived by the subscribers!

For handover, the picture is not so clear. If the handover is decided to salvage a rapidly degrading situation, queuing cannot help much since the connection will most probably be lost if the channel cannot be allocated right away; but if the connection is not in immediate jeopardy, and the handover decision just stems from general optimisation reasons, queuing can be a source of improvement. Thus, a correct use of queuing at handover requires that the handover process distinguishes several cases.

Fairness between the treatment of requests is to be sought when applying queuing. When no other consideration intervenes, the order of service is usually first-come, first-served: there we have the classic queue. More sophisticated schemes can also be devised, granting requests on the basis of their estimated priorities; handovers may be considered of higher

priority than re-assignments, and emergency calls are considered more important than other communications.

Even without queuing, there are ways to bias the granting of channel requests in congested situations. For instance, depending on the state of channel congestion, it could be worth rejecting the channel requests for outgoing calls to privilege those for incoming calls (a mobile user is more likely to answer than a fixed one if the mobile station is switched on, and the end-to-end circuit is almost entirely established already in the case of an incoming call). An even more drastic approach consists in forcefully terminating a connection estimated of a low importance, in order to reuse the resource for some needs deemed more important. This approach is referred to as pre-emption, and should be used with great care, given the negative impact on the preempted user.

For all the above reasons, some mechanisms have been introduced in the *Specifications* to enable priority strategies for radio channel allocation in the BSC. These mechanisms consist in conveying category, priority and pre-emption indicators, which can be used to influence the allocation decisions in cases of congestion. The *Specifications* do not describe how to use these indications, since the allocation schemes are left to the operators or manufacturers. Only their transport on interfaces is specified, and their actual use is indeed a source of difference between the products of different manufacturers.

Minimum indications relative to the purpose of the access are provided by the mobile station to the BSC at the beginning of the access procedure, to allow some priority mechanism for the initial allocation. Similarly, the MSC may provide some information when it requests a change of channel type (subsequent assignment procedure). In the case of handover the requesting BSC can also provide the MSC with some information about the reason for handover, and this reason can be carried up to the target BSC.

To summarise, a few things here and there in the procedures allow sophisticated allocation algorithms, with queuing, priorities, ..., but no constraints are put by the *Specifications* on the infrastructure equipment in this area. It is up to the manufacturers to include such functions, or for operators to request them.

Interference Considerations

The actual channel to be allocated may also be chosen with the optimisation of the transmission performance in mind. This requires the

BSC to have some knowledge beforehand about the performance of a connection for each of the free channels. Performance depends on many factors, most of them difficult to assess before the effective usage of the channel. However, there is one which is accessible: the level of interference in the uplink direction. When a channel is not allocated to any connection, the level received on the channel gives an idea of the level of interference and noise. GSM opens the possibility for the BTS to measure this reception level on all unallocated channels, and to transmit them regularly to the BSC. The BSC can then take this information into account in order to allocate a free channel of minimum uplink interference level, or to decide on an intra-cell handover if it is noticed that an active channel suffers a higher level of uplink interference level than the free one.

It should be noted that this feature is in most cases of secondary interest. The concern of an operator is to obtain a system able to support a maximum capacity; thus the situation to optimise is congestion. If all the channels of a cell are used at congestion (and this is the usual assumption), the allocation of the channels starting with the less interfered changes nothing to the eventual congestion situation. The only gain is the improvement of the average performance when far from the congestion state.

However, the assumption that all the channels are used at congestion does not always hold. One can imagine a cellular planning where the number of channels allocated to a cell is too high, in the sense that if all channels are used in all cells the overall quality of the connections is not acceptable. With such an approach congestion happens before all channels are allocated. A first consequence is that the BSCs must take care not to start additional connections when the congestion state is reached, even if channels are available. Another consequence is that taking uplink interference levels into account is then meaningful. What is obtained is a kind of automatic channel planning between the cells: the use of a channel in a cell will result in some interference level in other cells, thus preventing the latter to use interfered (and then interfering) channels. This is essentially a dynamic channel allocation method; it can be easily shown by taking the extreme assumption that all the channels are allocated to all cells. Without going to this extreme, a small over-dimensioning of the cell capacity plus this dynamic channel allocation method can be useful to cope with an abnormal distribution of the traffic between cells. If a cell is overloaded, but if the cells using interfering channels are not, the congested cell can with this approach use more channels than if all cells were equally overloaded.

Radio Channel Description

A side issue related to channel allocation is the way radio channels are described. The problem lies with frequency hopping. As explained in detail in Chapter 4, channel characteristics include time and frequency parameters. The time characteristics are not problematic; in the time domain, there are only 8 types of TACH/F and 68 types of TACH/8, depending on the type of channel and the offset relatively to the reference clock. One octet is therefore enough to code which of these 76 families of channels is being referenced (92 if half-rate channels were included).

In the frequency domain, things are somewhat more complicated. In the case of single-frequency channels, the number of different cases is 124 for GSM900 and 374 for DCS1800. In order to cope with some evolution, this frequency is coded on 10 bits, leading to a total of 18 bits to encode any single-frequency channel. However, when frequency hopping is being used, the number of combinations explodes. For GSM900 alone, a rough assessment shows that there are about

66,141,633,339,297,631,280,564,218,199,442,383,724,544

different possible hopping sequences compliant with the *Specifications* (including single frequency cases). This value needs 135 bits to be written in binary format. The matter is exponentially worse for DCS1800. This causes a problem because of the size of signalling messages, which is particularly critical on the PAGCH on which all initial assignment messages are sent. Some kind of reduction was therefore desirable. But which reduction?

First, what is needed to describe a hopping channel? The list of frequencies used by the channel, obviously. This frequency list, called the *MOBILE ALLOCATION* in the *Specifications*, is evidently the main source of length. The description also contains two other parameters used for the computation of the hopping sequence:

- the Mobile Allocation Index Offset (MAIO), of which there are as many possible values as there are frequencies in the list, hence its name since it describes the starting point for the hopping recurrence function; and
- the Hopping Sequence Number (HSN), which can assume 64 different values.

These two parameters fit on at most 13 bits.

The real coding problem lies with the coding of the frequency list. A first simplification consists in noting that only a set (i.e., an unordered

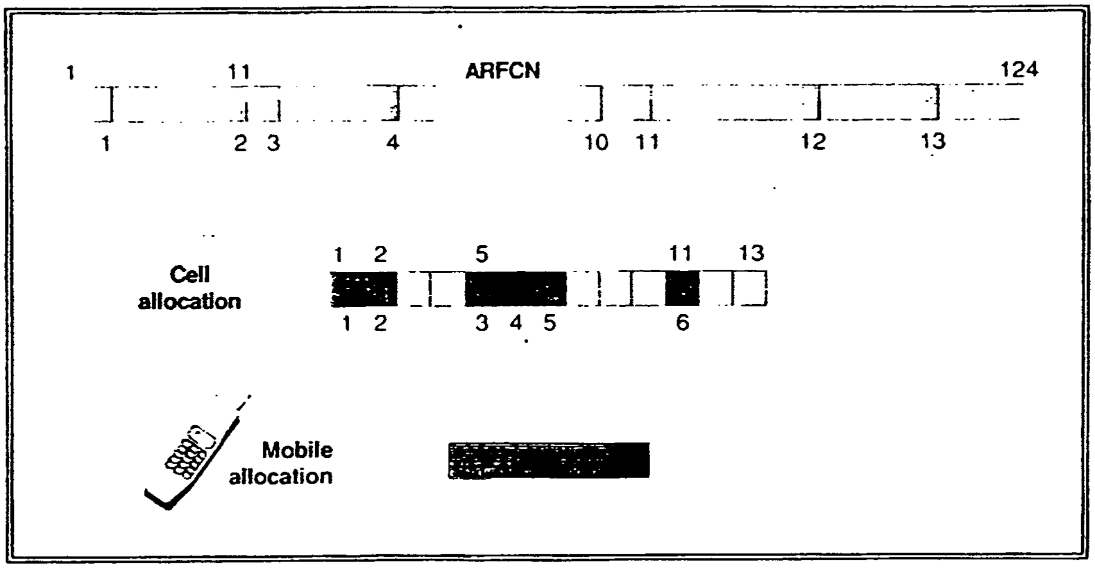


Figure 6.17 – Cell and Mobile Allocation

The cell allocation is the portion of the total resource usable in a cell, with regard to which each mobile allocation may be defined.

list) of frequencies is needed to define a channel. The hopping sequence generation algorithm refers to a list of frequencies, but the ordering is implicitly defined based on the respective value of the frequencies in Hertz. Therefore only sets are required, reducing the coding requirement to 124 bits. This is still too much compared with the constraints of the PAGCH.

Two approaches were possible: either restrict the number of possibilities, or design an efficient and sophisticated signalling scheme. Restricting the number of possibilities was inescapable for DCS1800. The maximum number of frequencies used by any given channel varies from 32 to 64 depending on the frequency range in which the channels are spread. But there is still with this limitation more frequency sequences in DCS1800 than in GSM900.

There remains the sophisticated signalling approach. All channels in a given cell use only those frequencies which are allocated to the cell by cellular planning. If this usually rather short list can be broadcast to mobile stations independently from any allocation, the description of a given channel in a known cell needs only to indicate which of these cell frequencies are being used. This "two-step" mechanism (see figure 6.17) leads to the concept of cell allocation, or cell channel description, which are the terms used in the *Specifications* to designate the set of all frequencies which are used in a given cell. (In fact, the main use of the cell allocation is for the initial assignment, so the cell allocation can be limited to the set of the frequencies used by channels that may be

allocated as initial channels.) The coding of the cell allocation uses 16 octets in GSM900 (where it includes a bit map for all 124 frequencies) and up to twice as much in DCS1800 (where the coding algorithm is much more sophisticated). This cell allocation is broadcast regularly on the BCCH. When the BSC sends a channel allocation message to the mobile station (whether at initial assignment, subsequent assignment or handover), it is able to encode efficiently the channel frequency list as a subset of the cell frequency list. If the latter includes n frequencies, the encoding of all possible subsets needs theoretically only n bits. The cell allocation in GSM900 is limited to less than 64 frequencies (since the bit map in the mobile allocation is limited to 64 bits). This is not really a constraint, since all frequencies cannot be used in all cells.

In the vast majority of cases, the cell allocation will not contain more than 32 frequencies; indeed, the gain brought by frequency hopping increases very little with higher numbers. Some encoding schemes have been introduced in the *Specifications* to encode efficiently the frequency set when the cell allocation consists in such a small number of frequencies. For historical reasons, two such schemes can be found: one for GSM900, one for DCS1800. The first scheme is reflected in what is called the *FREQUENCY CHANNEL SEQUENCE* element in the *Specifications*. This element allows to bypass the two-step approach described above, and is used at handover. Afterwards, another scheme was designed to cope with more frequencies, whilst maintaining compatibility with the first one. The latter scheme is indeed able to cope with 1024 different frequencies, leaving some room for extensions in the future of GSM-related systems...

6.2. ARCHITECTURE AND PROTOCOLS

The radio resourçe management functions are mainly dealt with by the BSS, and in particular by the BSC which acts as the orchestra conductor for these functions. It directs the mobile station as well as the infrastructure machines involved, which are more or less slave (in the RR field), i.e., the BTS (and TRAU) on one side, and the MSCs on the other. The BTS and the TRAU are indeed the main performers in the transmission chain, and they must as such be controlled by the Radio Resource functions, in fact by the BSC. The BSC is little involved in the transmission functions, but takes care of the consistency of the transmission chain, whether for its different properties (transmission

mode, cipher mode, ...) or for the quality of transmission (handover preparation and execution co-ordination).

If it is the anchor MSC which decides which properties of the transmission chain are desirable to fulfil the service, the role of the MSC for radio resource management is limited to some of the handover aspects. The anchor MSC is in charge of performing subsequent inter-MSC handovers when so decided by the serving BSC. The relay MSC is in charge of the handovers when between two cells of different BSCs under its control, and of the circuits between itself and its BSCs. For all other functions, the relay MSC acts as a transit node for signalling exchanges between the anchor MSC and the BSC or the mobile station. Functionally the relay MSC functions are all within the Radio Resource management realm.

This general description of the roles of the different nodes has to be refined for the handover function. Handover preparation makes use of a number of pieces of information, described in the above sections, and which originate from different sources. The BTS is the infrastructure entry point for all measurement information (both for reports from the mobile stations and for its own measurements). The BSC is the warden of all frequency planning and cell layout data. Information about traffic is spread between the BSC and the MSCs. Thus, whatever functional split is chosen, some real-time information transfer is needed between these entities in order to obtain a coherent handover strategy.

The basic split lies between the BSS (BSC + BTS) and the MSCs. The general rule puts the BSS in charge of the management of radio resources and of the decision to perform a handover on a given RR-session. The split was not clear from the start in the standardising process, and triggered many a discussion. The main problem between the MSCs and the BSS is taking traffic into account (either to influence cell choice, or for traffic handover). One could distinguish two approaches: either to have the MSCs indicate to its BSCs the level of traffic of surrounding BSCs, or to let the MSCs intervene in the handover algorithm to ponder radio criteria coming from the BSS with traffic considerations. The second approach was finally the one chosen, even though it may not be the simplest one. The intervention of the MSCs in the handover process blurs somewhat the functional border between BSC and MSCs, and their respective role is not easy to describe, especially for traffic handovers. This is nevertheless what we will attempt to do now.

First of all, when a BSC decides that an outgoing handover is necessary, it will indicate to the relay MSC one or more target cells, possibly managed by different BSCs and MSCs. Rescue handovers might call for several target cells, whereas confinement handovers obviously call for a single choice. If several targets are proposed, the relay MSC

may just try them one after the other, in the indicated order. Or it may choose among these ranked possibilities, taking into account its own traffic data. For a cell controlled by another MSC, the relay MSC forwards the request to the anchor MSC; at this level only one target cell is proposed at a time. The choice between several targets is then a function of the relay MSC not of the anchor MSC.

Another possibility for traffic handovers also exists in the *Specifications*, allowing the relay MSC to force the BSC to hand over a portion of the traffic from a cell, with the BSC being in charge of choosing which connections to hand over and to which cells; this is called the "candidate enquiry procedure". There also this possibility is not open to the anchor MSC.

In all cases, the conflict between confinement criteria and traffic criteria is obvious, and is not solved since data related to these criteria are under the control of separate nodes. In fact, there exists in the *Specifications* a means to group all data in a single place. It consists in transporting up to the relay MSC all the raw measurements coming from the mobile station and the BTS, and leave everything to the relay MSC for decision and choice. However, this possibility is not implemented by any manufacturer, and will disappear in phase 2, since the incurred load on signalling links and on the MSC processors would be formidable.

The functional split inside the BSS was also an area of long debates. The split can, there again, be done in different ways. One of the options is to group all the processing in the BSC, with all measurements being transmitted by the BTS without it performing any computation on them. The advantage of this solution is the centralisation of all data, so that handover decisions are taken based on the best data available. The drawbacks are the high signalling load on the Abis interface, since the incurred load of 2 messages per connection and per second represents by far the dominant traffic on this interface, and also an important requirement for high computational power in the BSCs. Another approach, referred to as "pre-processing" in the *Specifications*, consists in letting the BTS do an important part of the job, thereby relieving the Abis interface from the major part of its traffic and decentralising the computing load. Pre-processing is an accepted alternative in the *Specifications*, but the exact functional split between BTS and BSC in this case is not specified. Messages and information elements exist to cater for this function, but their semantics are left open to operators and manufacturers.

Pre-processing is often thought of as a way to reduce load on the Abis interface simply by making the BTS perform some averaging on the measurements. Such a simple approach could slow down the handover decision process significantly, which can be a source of inefficiency, in

particular in a small-cell environment. A correct usage of pre-processing introduces sophisticated algorithms in the BTS, including some part of the decision-making process. This is the main reason why details are not specified. It would indeed require too great a workload before a scheme could be shown and accepted by the specification committees.

All these different options related to the split of handover preparation functions within the infrastructure are taken into account in the *Specifications*. More details about the relevant procedures can be found in the SACCH procedural section, as well as in the handover execution section.

Protocols

Independently of the infrastructure architecture, the implementation of the RR functions requires some kind of protocol between the mobile station and the network. On the network side, the interlocutor (or peer entity) of the mobile station for this protocol is the

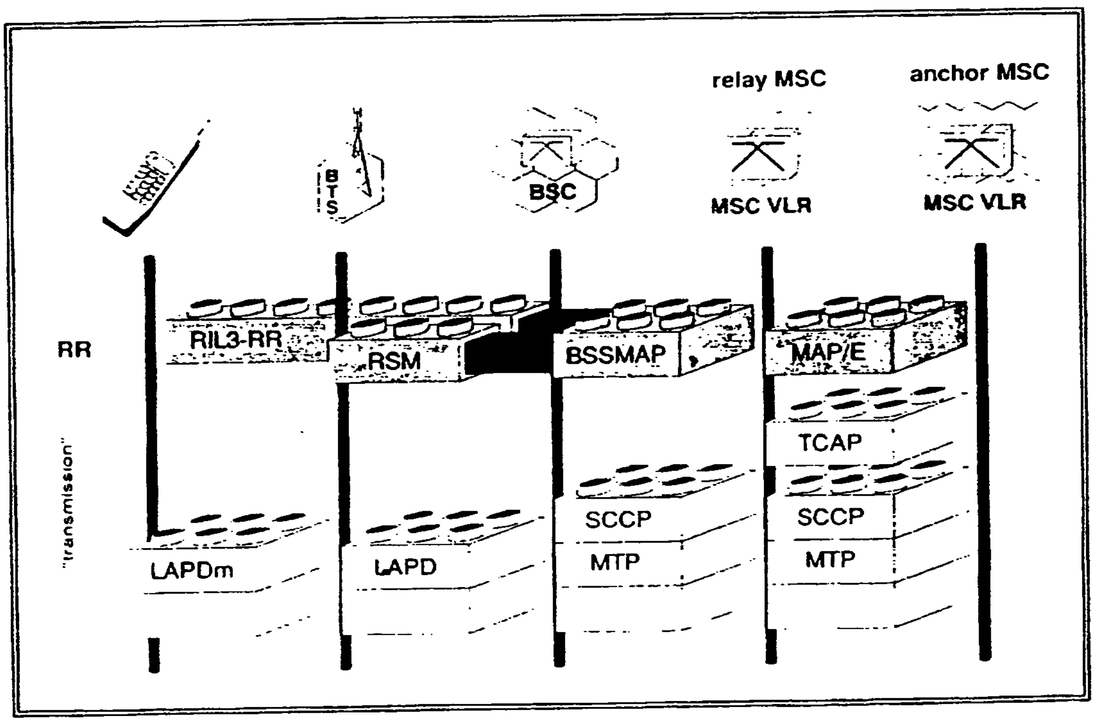


Figure 6.18 – RR protocol architecture

Protocols for RR management are needed on many interfaces, including the A interface (BSSMAP), Abis interface (RSM) and the MSC-MSC interface (MAP/E).
The main co-ordinator is the BSC,
which is also in relation with the mobile station (RIL3-RR)

BSC (in fact, a small part of the signalling is also handled by the BTS for efficiency reasons). This protocol will be denoted RIL3-RR.

The functional distribution between infrastructure entities calls for other protocols on terrestrial links: one between BTS and BSC, one between BSC and relay MSC, and one between relay MSC and anchor MSC. The first one, on the Abis interface, is used for the BSC to configure the transmission path and for the BTS to report measurements to the BSC. It has no official name in the *Specifications* (experts refer simply to the 08.58 protocol, from the number of the corresponding Technical Specification), and will be here referred to as RSM (Radio Subsystem Management).

The protocol between BSC and relay MSC, on the A interface, is used to carry the requests for initial connection establishment, as well as for any change in the connection attributes according to upper layer requirements. It is used also for handling handovers between the relay MSC and the BSC. This BSC-MSC protocol is called the BSSMAP protocol (BSS Management Application Part).

The last protocol, between two MSCs of adjacent coverage areas and supporting the exchanges between relay MSC and anchor MSC, is part of the MAP and will be referenced MAP/E protocol. Figure 6.18 shows the machines involved in radio resource management and the protocols between them.

6.3. RR PROCEDURES

In the first part of this chapter, we looked at the different tasks needed for the management of the radio resources from an "object" point of view. We have studied how for instance the channels are managed, or how timing advance is controlled, but all these topics were seen rather independently. In this second part, we will look at the details of what happens at different moments, combining all these independent aspects. Most procedures in the RR area are concerned with several functional aspects simultaneously, and involve, or may involve, all machines between the mobile station and the MSC. We will then revisit a number of the topics we have seen in the previous sections, but with stress on the temporal relationships.

We will start at the beginning, that is to say with the access procedure, where the RR session is created. This will be the occasion to look in detail at the use of the RACH for the initial contact from the

mobile station. Next we will present the paging procedure, which precedes the access when the requirement for a connection comes from the infrastructure side. The following sections will be devoted to what happens during the life of an RR session. One aspect concerns the change of some characteristics such as the type of channel or the ciphering mode, when requested by the anchor MSC. Another aspect discussed is the execution of a handover, with all its variants, including call re-establishment, presented here as a last-resort kind of handover. The study of the main adventures of an RR-session will end with the release procedure.

We will then discuss a number of ill-assorted procedures, such as the procedural handling of the signal measurements, the timing advance and the transmission power control, which is done on the SACCH; the frequency redefinition procedures, which are rather complex mechanisms to cope with a change in the allocated frequencies in a cell when frequency hopping is used; and finally the broadcasting of various information on the BCCH.

In all cases, in this chapter and the following ones, what normally happens will be presented with some details, without going, however, into the internal structures of the signalling messages. The rarer cases, involving failures or collisions between events, will at best be hinted at, though they are maybe the most important aspects to have in mind when designing signalling protocols. But even a minimum attempt to correctly cover this subject would need a far larger text than can be included here.

6.3.1. INITIAL PROCEDURES: ACCESS AND INITIAL ASSIGNMENT

The purpose of this section is to describe the procedures enabling the transition between the two major states of a mobile station, i.e., "idle" mode (where the mobile station is everything but idle, but refrains from any active transmission towards the infrastructure), to "dedicated" mode where the mobile station is actively transmitting on a channel allocated for its own use. The transition corresponds to the establishment of an RR-session (see page 313).

The initial assignment procedure is always triggered upon the request of the mobile station, for one of three major reasons:

- to perform location updating;
- to answer to a paging; or

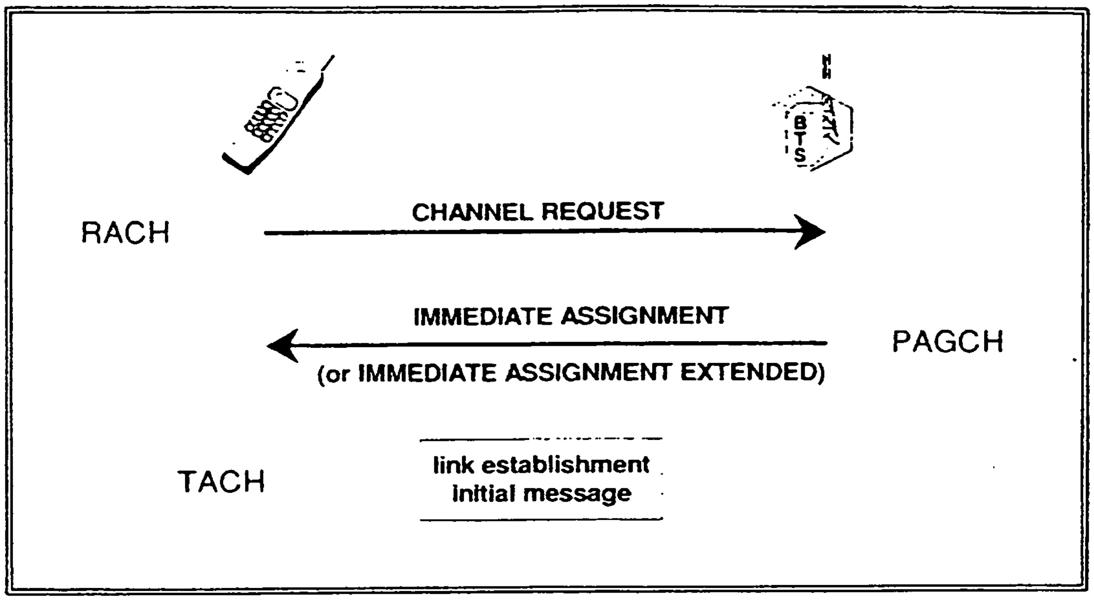


Figure 6.19 – Initial access procedure

The transition from "idle" to "dedicated" mode is always triggered by the MS, through an RIL3-RR CHANNEL REQUEST message sent on the random access channel. Only when the signalling link layer has been established and an "initial message" sent on the new dedicated channel does the network know the identity of the MS.

- as a result of a user's request, i.e., for an outgoing call, a supplementary service management request, or the sending of a short message.

In all cases, the access procedure is the same (see figure 6.19) In broad terms, this procedure starts with an RIL3-RR CHANNEL REQUEST message sent on the RACH; the answer from the network is conveyed in an RIL3-RR IMMEDIATE ASSIGNMENT (or RIL3-RR IMMEDIATE ASSIGNMENT EXTENDED) message sent on the Paging and Access Grant Channel (PAGCH), conveying the description of the channel allocated to the mobile station; finally, the mobile station establishes the link layer for the transfer of signalling on the newly allocated channel, and sends a first signalling message on this channel (the "initial message"), conveying the subscriber's identity and the reason why it requests a connection. This basic canvas appears very simple at first view; but each of the corresponding steps reveals some complexity when studied more closely.

6.3.1.1. Random Access

The channel request message is a curious animal indeed, and deserves some attention. The network has no method of knowing when

mobile stations will need to communicate, and therefore this first message from the mobile station cannot be scheduled to avoid the simultaneous transmission of more than one mobile station (a collision). This is the major problem of random access schemes, and the channel name (Random Access CHannel, or RACH) indeed expresses the fact that mobile stations transmit independently from one another.

Of course, mobile stations do not transmit at any time, but follow the slotting of time imposed by the TDMA scheme (a RACH/F uses only one slot every 8 burst periods). Collisions may therefore be studied slot by slot. When two mobile stations transmit during the same slot, two things may happen: either one of the bursts is received by the BTS at a level significantly higher than the other one, allowing its correct decoding (this is called a "capture"), or none is received correctly. Collisions are therefore a source of message loss, which increase with traffic. In order to provide a satisfying rate of success for access attempts, repetitions must be used. The repetition scheme cannot be too simple, otherwise its effect on the throughput in a high load situation can be disastrous, and may indeed lead to a complete deadlock situation. This kind of problem has been thoroughly studied in the field of random access techniques, an area of interest for many networks using shared resources (Local Area Networks, Packet Radio, ...). GSM offers an example of one of the best-known (and simplest) random access schemes, with the RACH being an application of the so-called "slotted Aloha" protocol.

When a request has not been answered, the mobile station will repeat it. If two mobile stations whose attempts have collided would choose to repeat them some given constant time afterwards, their requests would collide again. Repetitions on the RACH are therefore performed after a "random" interval to avoid this phenomenon. As in all Aloha protocols, this re-transmission strategy is not enough to escape from collapse when the traffic goes over a given threshold. In practice, the offered load (in terms of number of requests, which can lead to one or more messages sent) on the RACH should not exceed about a quarter of the total sending opportunities (number of slots on the channel). In order to control this load, GSM uses three different means, corresponding to three different kinds of overload.

The collapsing threshold depends on the number of repetitions and the average time between them; one way to make the RACH robust to a higher load is to spread these repetitions further apart, and/or to reduce the number of repetitions. Of course, this may be detrimental to the quality of service, respectively in terms of delay or in terms of success probability. Such methods cannot be pushed very far and are not sufficient to control very high loads. It is nevertheless useful in cases of small and temporary overload. In GSM, both the number of repetitions

parameter	resulting value
TX-INTEGER	random scheduling of each attempt over 3 to 50 slots
MAX RETRANS	up to 1, 2, 4 or 7 repetitions allowed

Table 6.3 – RACH repetitions control parameters

Both the time interval between re transmissions of a random request and the maximum number of such repetitions are controlled by parameters.

and the intervals between them are controlled through parameters broadcast regularly on the BCCH on a cell-per-cell basis. Since it is important to have a short delay between the moment when the BSC decides to change these parameters and the moment when the mobile stations act on them, it has been decided to send them in all BCCH messages, i.e., 4 times per second. The scheme is controlled by two broadcast parameters, the average time between repetitions (*TX-INTEGER*), and the maximum number of allowed repetitions (*MAX RETRANS*, see table 6.3).

These parameters should be controlled through a feedback loop taking into account the observed throughput. It should be noted that Aloha control is in fact not absolutely necessary: the values can be set to a constant choice representing some compromise between throughput and delay without jeopardising the system. As already mentioned, this mode of control can only cope with brief (in the order of one second) traffic peaks or sustained marginal overload.

Before studying further the other means to control the load on the RACH, it is worth noticing that this channel constitutes only the first link in a chain of resources, and is not necessarily the bottleneck of the system in congested situations. There are indeed other candidates for such a role: the PAGCH, which offers a limited capacity for carrying both initial assignment and paging messages, is one of them. Channel allocation is another, since the pool of available channels is also limited in each cell. An efficient overload control takes all these factors into account and tries to cut the traffic at the source (i.e., on the RACH) in cases of congestion. This means that overload control mechanisms must not necessarily try to maximise the RACH throughput, but must limit it to the maximum traffic the whole chain can swallow.

Having said that, the second way of "controlling" the load on the RACH consists of rejecting the requests with a message forbidding the mobile station to access the channel for some specified length of time. This mechanism prevents any further repetitions by the mobile station, either through its automatic repetition scheme (controlled as explained above), or through repeated user requests, a natural tendency of users

upon failure. Obviously, this mechanism should only be used when overload threatens the overall throughput, since it significantly increases the service time. Besides, it requires the sending of a reject message (the RIL3-RR IMMEDIATE ASSIGNMENT REJECT message) from the BSC to the requesting mobile station, and is therefore inappropriate when congestion impacts the Abis interface or the PAGCH.

Last of all, a third (and most robust) line of defence exists and makes use of the concept of access class. Basically, it consists in forbidding whole populations of mobile stations to access the cell, through an indication on the BCCH. This scheme is very efficient, since it enables a cut-down in the traffic at the very source, without incurring any additional traffic towards mobile stations. In order to achieve this, subscribers are split into 10 balanced sub-populations, through a random allocation controlled by their home PLMN operator. The access class to which a given subscriber belongs is stored in the SIM and is therefore available to the mobile station. In normal load situations, all classes are allowed access. When traffic must be cut down, the BSC can decide to block 1, 2 or any number of these access classes, reducing statistically the amount of traffic by 10%, 20%, ... Mobile stations belonging to the forbidden classes refrain from accessing the network, except in specific cases (e.g., emergency calls, which are controlled by a specific indicator). In order to be fair, if the overload period lasts for a fairly long time, the BSC must take care to change the set of authorised classes regularly (though this must be handled with care, as for instance all the mobile stations waiting for a location updating will try to access at the very moment their class is authorised).

To avoid blocking special categories of users in congested situations, five more classes are defined, for "Very Important GSM Subscribers", as shown in table 6.4. Access for these classes is also controlled through indicators broadcast on the BCCH. The corresponding

"special" access class	subscriber category
11	left open to the PLMN operator
12	security services
13	public utilities
14	emergency services
15	PLMN staff

Table 6.4 – Access classes for "Very Important GSM Subscribers"

In addition to the "standard" random subscriber classes (0 to 9), the SIMs of specific users may feature one of the above "privileged" classes.

"privileged" subscribers belong both to one of the 10 standard classes and to one (or several!) of the special classes, and may access the network when at least one of their classes is allowed.

One reason for access escapes this class-dependent mechanism: emergency calls. In order to be able to control this source of traffic, the BCCH indicates instead whether emergency calls are allowed or not, and this is applied to all subscribers.

In a situation of congestion, one may wonder whether a mobile station whose access is either unauthorised or rejected in a given cell, should be allowed to try to access through other cells. There are pros and cons in this approach. When traffic is spread unevenly in an area, the overload situation of a cell can be improved if part of the traffic is redirected to neighbour cells. However, as explained in the section dealing with handover preparation, there exists one best cell for each mobile station for optimising the general interference level, and the choice of another cell contributes to a degradation in terms of spectral efficiency, and such a strategy will spread the congestion. But if the congestion situation is general, to allow the mobile station to try elsewhere would simply tend to group the mobile stations of given access classes in given cells, which would diminish if not negate the control effect, and would not be the best for interference, to say the least.

The *Specifications* distinguish different cases for this issue. When the mobile station has tried to access and has failed, it is allowed to select another cell and is indeed forbidden to choose its previously serving cell for at least 5 seconds. This enables the mobile station to attempt an access on the second best cell (at least if the user so wishes, or automatically). Because of this mechanism, some of the traffic will be diverted to neighbour cells in situations of local congestion or when the number of repetitions is reduced. In other cases, i.e., upon explicit access rejection or when the relevant access class is barred, the *Specifications* require the mobile station to stay in or leave the serving cell in the normal way, preventing an increase in congestion. This is at the expense of a "no service" period for some mobile stations, but the average service per mobile station is in fact increased through such a "drastic" mechanism. As already mentioned, fairness can be obtained by rotating the blocked classes.

6.3.1.2. The Contents of the RIL3-RR CHANNEL REQUEST Message

The RIL3-RR CHANNEL REQUEST message, sent on the RACH, is very short indeed. Its useful signalling information consists of just 8 bits!

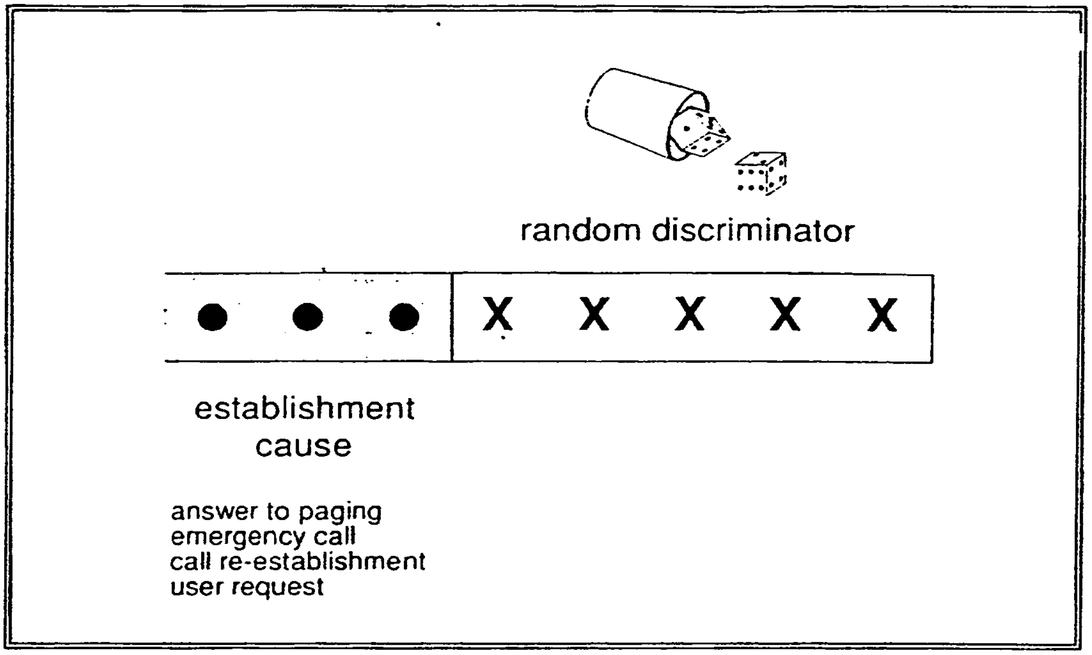


Figure 6.20 – Useful contents of an access burst on the RACH

Only 8 bits are available in this short type of burst.
Three of them indicate the reason for access, and the remaining five serve as a random discriminator.

This capacity is obviously insufficient to carry all the information the mobile station would want to transmit, such as the subscriber's identity, the reason for requesting a channel, the characteristics of the mobile equipment, ... All of this information is in fact included in the "initial message" which will be the first information transmitted on the dedicated channel, once allocated.

But the most critical usage of a discrimination between random access attempts is not to provide information to the network. A given mobile station must be able to correlate an initial assignment from the network with its own request, with as little ambiguity as can be achieved. For this purpose, 5 bits among the total of 8 are chosen randomly by the mobile station, reducing drastically the probability that two mobile stations send identical messages during the same slot, which may in case of capture lead to an ambiguity as to which of the two requests is being granted.

Three bits remain, which are used as shown in figure 6.20 to provide a minimum indication of the reason for accessing the network. This first rough indication may be useful for discriminating rejections in case of congestion, and also to choose the best type of channel to allocate.

6.3.1.3. The Initial Channel Assignment

After the BTS has correctly decoded a channel request, it indicates it to the BSC through an RSM CHANNEL REQUIRED message, with one important piece of additional information: an estimate of the transmission delay (this indication is critical to initialise the timing advance control). A field of unspecified content (*PHYSICAL INFORMATION*) allows the manufacturer to add more information, such as the reception level.

In normal load situations, the BSC then chooses a free channel (TACH/8 or TACH/F), activates it in the BTS, and, when the BTS has acknowledged this activation, builds an initial assignment message to be sent on the PAGCH.

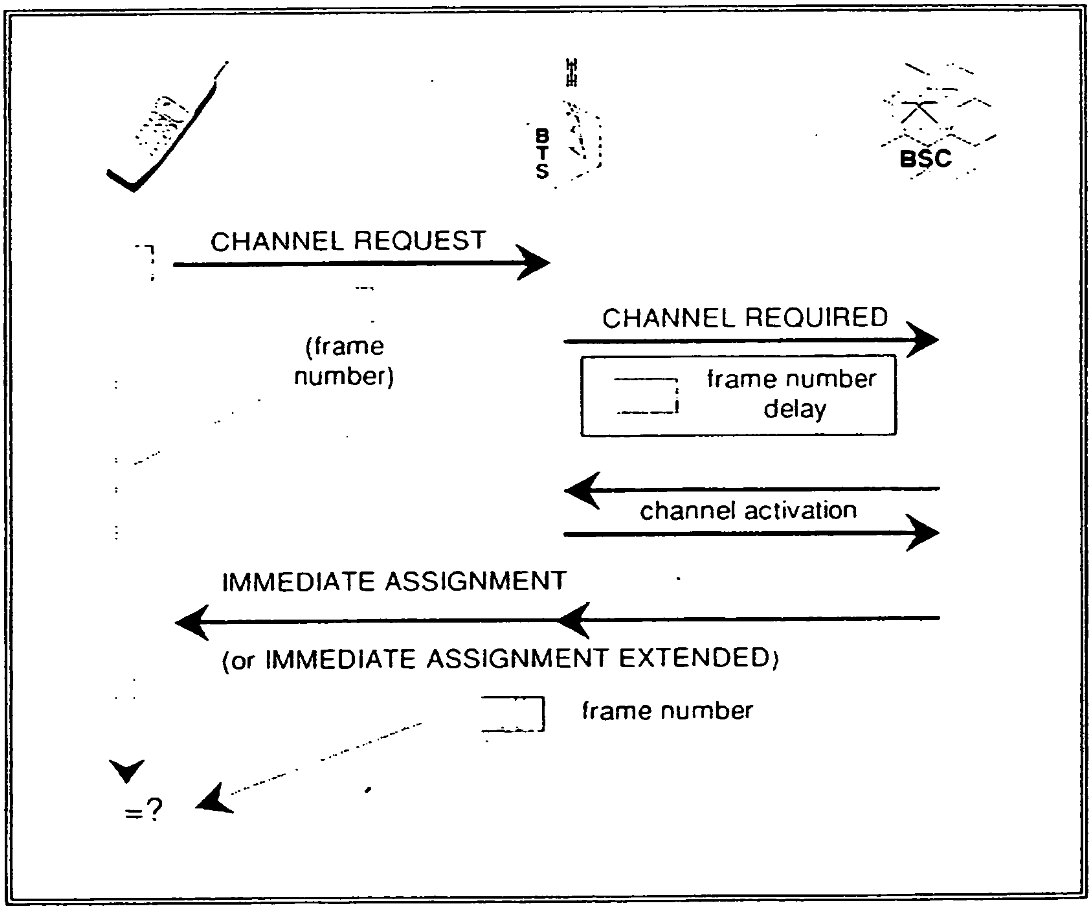


Figure 6.21 – Initial assignment procedure

After the activation handshake on the Abis interface, the BSC prepares an initial assignment indication containing the 8 bit discriminator as received in the correctly decoded RIL3-RR CHANNEL REQUEST message, as well as the frame number in which it was received. This enables the mobile station to check whether it is concerned with the message or not.

The activation process requires the BTS to prepare for the access of the mobile station on the newly allocated channel. The timing advance is initialised based on the transmission delay estimate which the BSC indicates (back!) to the BTS. Even though this estimate has been initially calculated by the BTS, this going back and forth between BTS and BSC is necessary, since the BTS has no means to correlate the messages received on the RACH with corresponding channel assignments.

The initial assignment indication sent to the mobile station on the PAGCH contains the description of the allocated channel, the initial timing advance to be applied, the initial maximum transmission power, as well as a reference allowing all the mobile stations expecting such a message to know whether they are being addressed or not. This last point is worth some more explanation.

Addressing is done by including in the initial assignment indication the exact contents of the RIL3-RR CHANNEL REQUEST message which is being answered, plus the time reference of the slot in which it was received (such a time reference exists thanks to TDMA). This allows mobile stations to check whether they are actually concerned by each initial assignment, by comparing these values with the ones they have stored when sending the RIL3-RR CHANNEL REQUEST message, as shown in figure 6.21.

Besides, answers to RIL3-RR CHANNEL REQUEST messages can be sent in any block of the PAGCH, even on the paging sub-channels. As a consequence, once a mobile station has made an access attempt, it should monitor the whole PAGCH (of the same Timeslot Number as the RACH it used for access) for an answer from the network. Furthermore, the BCCH messages must be decoded continuously during this period, in order for the mobile station to set the RACH control parameter values in real-time. This phase is very constraining for mobile stations in terms of reception (40 bursts every 51×8 burst periods), almost comparable to TACH/F reception.

Let us examine some side issues. It may happen that the reaction of the infrastructure to an RIL3-RR CHANNEL REQUEST message is too slow to avoid a repetition from the mobile station. In such inefficient situations a given mobile station may be allocated a channel twice (or even more times), since the infrastructure has no means of knowing whether an RIL3-RR CHANNEL REQUEST is the repetition of a previous one or not. The mobile station will use the channel allocated in the first initial assignment message it decodes, and the other ones will have been blocked for a few seconds in vain. Nevertheless, the *Specifications* require that the mobile station be able to accept the network answer to any of its last three RIL3-RR CHANNEL REQUEST messages, in order to get service from such not-so-efficient BSS equipment.

In congested situations, when no channel is free for allocation, the BSC may choose not to answer to an RIL3-RR CHANNEL REQUEST message, or to send back a rejection indication. The first choice is not very efficient, since the mobile station will repeat its attempt. Explicit rejection is done through an RIL3-RR IMMEDIATE ASSIGNMENT

REJECT message, containing a time indication during which the mobile station is forbidden to make any more attempts on the RACH (the *WAIT INDICATION* parameter). If the overload situation does not concern the RACH, the value of the *WAIT INDICATION* can be null, otherwise it is a useful mechanism to help reduce RACH load (see page 370).

The PAGCH is an important potential bottleneck of the system. In order to improve its efficiency, both initial assignment indications and channel request rejection indications can be grouped together to form messages. There are two assignments in an RIL3-RR IMMEDIATE ASSIGNMENT EXTENDED message, and up to four rejections in an RIL3-RR IMMEDIATE ASSIGNMENT REJECT message. Though this point is unclear in the *Specifications*, the original intention was for the BTS to perform the grouping on the basis of the individual indications provided by the BSC. The BSC has the possibility to provide the BTS with immediate assignment indications which are not ready-to-send messages (using the RSM IMMEDIATE ASSIGN COMMAND message). The BTS must then build the corresponding RIL3-RR messages The fact is that the BSC *can* also perform the grouping and build the messages, and can provide ready-to-send messages to the BTS (this is, by the way, the only possibility for assignment rejection messages). However, nothing really precludes the BTS to un-build these messages and to build others grouping the requests differently. The same debate exists for the paging indications.

In no case does the BSC schedule the transmission of the assignment indications: this is much easier for the BTS. The counterpart is that, despite grouping, congestion may happen. This is resolved simply by the BTS, which drops messages which are in excess of the achievable throughput. While this resolves the BTS congestion problem, for the BSC it results in TACHs allocated but not "assigned". To avoid a worsening of the congestion situation through this effect, the BTS can indicate the non-sending of a message with an RSM DELETE INDICATION message. A second effect of the message is to indicate the overload to the BSC.

6.3.1.4. The Initial Message

Once it has received an initial assignment indication, the addressed mobile station modifies its reception and transmission configuration to adapt it to the frequency and time characteristics of the new channel. In phase 1, this new channel can be a TACH/8 or a TACH/F, always in "signalling only" mode (see page 321). The transmission level is set to a value broadcast on the BCCH (or to the maximum transmission level of the mobile station, whichever the smaller), and the transmission starts with the timing advance value specified by the BSC.

The first thing the mobile station does on the new channel is to transmit a link layer SABM frame for SAPI 0, i.e., the frame used to establish in acknowledged mode the link layer connection for signalling messages.

In standard HDLC protocols, an SABM frame does not carry any information other than the one necessary for the link layer level. In GSM, the SABM frame sent within the initial access procedure contains a signalling message, the "initial message". The reasons for departing from standard usage are twofold. The first reason is efficiency, though this was

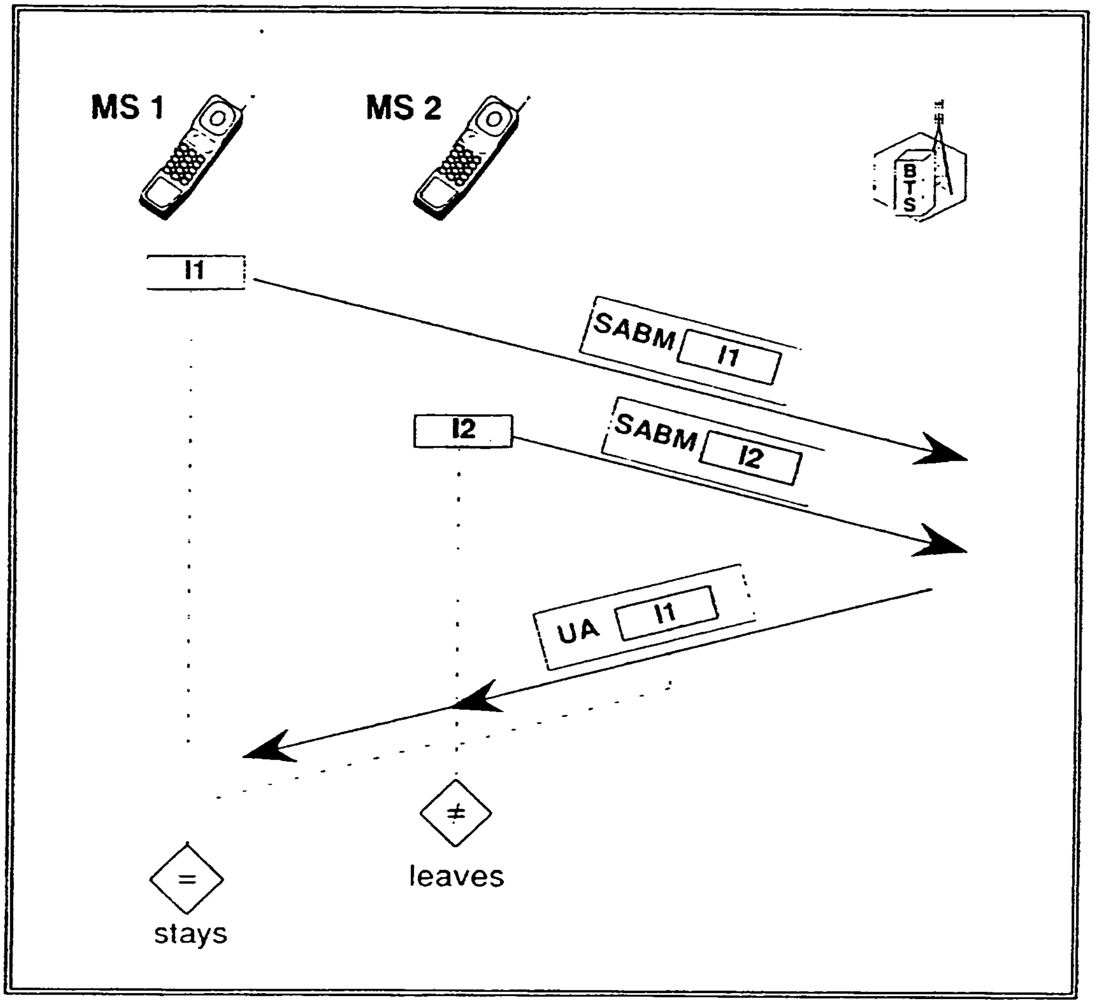


Figure 6.22 – Contention resolution at link establishment

In rare cases, more than one mobile station may find itself
on the same dedicated channel.
The transfer of a non-ambiguous initial message in the SABM-UA exchange
allows each mobile station to know whether the channel is for its own use or not.

not the leading cause. The other reason comes from the fact that the reference used in the initial assignment to address the mobile station is not fully unambiguous. It may indeed happen (though this is a rare case, around 1% in high load situations) that two mobile stations simultaneously send RIL3-RR CHANNEL REQUEST messages with exactly the same contents, and that one of them is correctly received and answered by the BSS. The ensuing channel assignment will be understood by both mobile stations as their own, and both mobile stations will access the "dedicated" channel. Therefore, until the point where mobile stations identify themselves in a non-ambiguous way, there is not a 100% guarantee that a single mobile station will access the channel.

As a consequence of this situation, potential collisions must be detected as soon as possible on the new channel, and the SABM-UA exchange provides this facility by including ("piggybacking") unequivocal information on these link layer frames. Mobile stations check the piggybacked contents of the UA frame. If one mobile station receives a UA containing something different from the contents of the SABM it sent, it must leave the channel and start the access procedure all over again, thereby enabling the "right" mobile station to stay undisturbed on its own channel (see figure 6.22).

One obvious way to obtain an unambiguous SABM content is to use an identity unique to the mobile station. This would be enough to serve the purpose of collision detection. However, the efficiency criterion was also taken into account and the SABM includes more than just this identity, and includes a full "initial message".

The "initial message" comes in four different brands, depending on the reason why the access was triggered (see table 6.5). All these messages contain an identity of the mobile station; the **classmark**, a field indicating some key characteristics of the mobile equipment, including the maximum transmission power; and complementary information making the reason for access more precise when need be. All but the first of those messages belong to the RIL3-MM protocol, and will be explained in Chapter 7. The first belongs to the RIL3-RR protocol, but could have as well been put in RIL3-MM. If there had been enough room, these messages would have been properly formatted, with a part for the RIL3-RR protocol (including the classmark), and the rest in an RIL3-MM part.

Reason for access	initial message
Response to a paging	RIL3-RR PAGING RESPONSE
Normal location updating, periodic location updating, "IMSI attach"	RIL3-MM LOCATION UPDATING REQUEST
IMSI detach	RIL3-MM IMSI DETACH
All other cases (call set-up, short message transmission, supplementary service management, ...)	RIL3-MM CM SERVICE REQUEST

Table 6.5 – Possible initial messages

Four different types of signalling messages may be used as the "initial message", depending on the reason having led to a channel request.

There is usually but one choice for this "initial message". One case of collision may arise, when, after having sent an RIL3-RR CHANNEL REQUEST message, the mobile station receives a paging indication and then the answer to its RIL3-RR CHANNEL REQUEST. What should the initial message be in such a case? Should it be consistent with the reason for sending the original RIL3-RR CHANNEL REQUEST message, or should it be the RIL3-RR PAGING RESPONSE message? It would seem fair to choose the RIL3-RR PAGING RESPONSE message if for instance the request from the mobile station concerned periodic location updating, but no priority scheme is defined in the *Specifications*, and the whole matter is left open for mobile station manufacturers.

Once an "initial message" has been received by the BTS (and sent back without any modification inside the UA frame), it is passed to the BSC in an RSM ESTABLISH INDICATION message. At this point, the mobile station classmark is stored for further use (e.g., to choose the power control loop parameters), and the BSC then sets up an SCCP connection towards the MSC (see Chapter 5). This is done through an SCCP CONNECTION REQUEST message, on which the initial message may optionally be piggybacked. Only then does the MSC become aware of the contact with the mobile station. The initial message, whether piggybacked or sent after the SCCP connection establishment, is carried in a BSSMAP COMPLETE LAYER 3 INFORMATION message, independently from its protocol (RR or MM). It contains enough information for the MSC to trigger required actions in the upper layers (MM, CC, ...), but this comes out of the scope of the access procedure.

When the access procedure ends, the RR-session is fully established with a complete signalling path between the mobile station and the MSC. With the establishment of the SCCP connection, the MSC takes control of the decisions concerning the transmission characteristics of the RR-session, and the BSS is at the ready whilst monitoring the transmission and performing handover decisions.

6.3.1.5. The Mobile Station Classmark

Mobile stations differ by many characteristics, such as their maximum transmission power and the services they may support. It is important for the infrastructure to be aware of some of these characteristics when the mobile station is engaged in a connection. Because the equipment of the user may be changed without warning the operator (subscription is linked to the SIM, not to the mobile equipment), this indication must be given at the beginning of each new connection. This is the purpose of the mobile station classmark. The full contents of the classmark are shown in table 6.6.

There exists in fact, for efficiency reasons, a subset of the full classmark, called "mobile station classmark type 1" in the *Specifications* (the full classmark being referred to as "mobile station classmark type 2"), destined to be sent in the RIL3-MM LOCATION UPDATING REQUEST message, so that this message fits in a single block (segmenting is not

parameter contained in the classmark
revision level
RF power capability
encryption algorithm
frequency capability
short message capability

Table 6.6 – Full contents of the mobile station classmark

The classmark identifies those characteristics of the mobile equipment which are needed by the infrastructure during a connection.

used when the message is piggy-backed). The shortened classmark is also used in the RIL3-MM IMSI DETACH message, for no obvious reasons.

The **revision level** is used for upward compatibility handling between successive phases of the *Specifications*. Mobile stations developed and type approved according to the phase 1 *Specifications* should set this value to 000. In the future, other values will be allocated, so that the infrastructure will know which level of upgrade is used by each mobile station.

The **RF power capability**, often referred to as the transmission power class, or even as the class, refers to the maximum power the mobile station is able to transmit. This information is used for power control and handover preparation. Power classes are not defined in the same way for GSM900 and DCS1800, as shown in table 6.7. Mobile stations of class 1 for GSM900 may possibly never be developed; indeed the typical classes in GSM900 are class 2 for portable or vehicle-mounted equipment, and class 4 for handheld. Class 5 handheld stations will most certainly experience strong limitations in coverage and will be destined for urban areas. The typical class of DCS1800 mobiles is class 1.

The **encryption algorithm** indicates which ciphering algorithm (if any) is implemented in the mobile station. In phase 1, there is but one choice: all mobile stations must implement the A5 algorithm specified by the GSM MoU. This field will enable the BSC to cope with future mobile stations with other ciphering capabilities, by being able to choose whether to cipher or not, and if yes with which algorithm.

The **frequency capability** is also present for future use. It enables the network to cope with mobile stations having different capabilities in terms of frequency bands. A phase 1 GSM900 mobile station must be able to cope with frequencies anywhere in the 2×25 MHz band allocated

Class	GSM900	DCS1800
1	20 W	1 W
2	8 W	0.25 W
3	5 W	–
4	2 W	–
5	0.8 W	–

Table 6.7 – Mobile station power classes

Five classes are defined for GSM900, and two for DCS1800, corresponding to the maximum transmission power of the mobile station.

from the start to GSM. An extension of this band to, say, 35 MHz is under study, but all mobile stations will not be able to cope with the extension. Therefore, the classmark will enable the BSC to distinguish the different populations of mobile stations and allocate channels to each mobile station according to its own frequency capability.

Last of all, the **short message capability** indicates whether the mobile equipment is able to deal with short messages. Such an indication is not strictly necessary in the classmark, since a mobile station not equipped to deal with short messages (if any!) can always inform the network by rejecting a LAPDm SABM frame on SAPI 3. But it is more efficient to know the capability of the mobile station from the start, since it enables the network to eschew the transmission of short messages between the short message service centre and the VMSC.

We have seen that the classmark is sent by the mobile station in the initial message, at the beginning of the RR-session. The network does not store the classmark between RR-sessions, since the user can change his equipment. Now, it may happen that the classmark changes *during* the RR-session. This is not a common event, but an example is a mobile equipment composed of a handheld part and a vehicle-mounted part including an RF transmitter, with the possibility to connect and disconnect the two parts during a communication. Then the power class changes, and the new value must be provided to the network. To achieve this, a procedure is included in the RR plane. The indication from the mobile station is carried in an RIL3-RR CLASSMARK CHANGE message and the BSC can forward the indication to the relay MSC with a BSSMAP CLASSMARK UPDATE message. If the relay MSC is not the anchor MSC, the chain stops there in phase 1, since the MAP/E protocol does not support this functionality.

6.3.2. PAGING PROCEDURES

A little has to be said on the paging procedures. When a call to a subscriber reaches a MSC through which the subscriber is deemed reachable, the MSC determines the location area where the mobile station is registered and sends a BSSMAP PAGING message to all the BSCs controlling cells in this location area. The message contains the subscriber identity to page with (it could be a temporary mobile station identity, called the TMSI or the full international mobile subscriber identity, the IMSI), the IMSI to determine the paging sub-channel (to cope with discontinuous reception), and the list of cells in which the paging must be issued.

The BSC in turn sends an RSM PAGING COMMAND to the BTS device in charge of the PAGCH of suitable TN (determined by the BSC, from the IMSI and the common channel configuration), for each cell in the list. This message contains the number of the paging sub-channel, which is computed by the BSC, as well as the TN of the PAGCH.

The BTS in turn possibly packs some paging requests together and sends the resulting messages on the correct paging sub-channel. As indicated in the functional section, more sophisticated approaches are possible.

Paging messages come in three brands (called RIL3-RR PAGING REQUEST TYPE 1, TYPE 2 and TYPE 3), adapted to the size of the identity used for paging. Type 1 can carry two identities of whichever sort, whereas type 3 can carry four TMSIs, and type 2 two TMSIs and one identity of whichever sort.

A topic of interest is the repetition policy. In other areas, GSM is specified so as to provide a correct quality of service when the transmission quality is far from perfect. To be consistent, the paging indication should not be sent only once in each cell. A typical value (not specified) would be to send it three times. No repetition mechanism is described in the *Specifications*, and this leaves some obscurity as to which machine takes care of the repetition: the MSC, the BSC or the BTS? To choose the BTS has some advantages: it can optimise the use of the PAGCH, and in particular it may repeat paging requests more than the minimum required if the channel has some room left. On the other hand, if the MSC is in charge of managing the repetition process, it will wait some time before requesting repetitions, thus avoiding useless repetitions in many cells in cases of successful mobile station access. Neither the BTS nor the BSC are capable of such monitoring, since they are not able to relate the mobile station answer to the paging. This seems obvious when the mobile station has accessed another cell (the MSC does not indicate it back to the BSCs from which it requested paging, it would be too big a procedure), but the same is also true when the mobile station

accesses through a cell under their control. A sensible scheme seems to provide two levels of repetitions, one in the MSC, with a long period, to cope for instance with short reception interruption (such as a change of cell, or the mobile station passing through a tunnel); and a short term repetition in the BTS, when load allows, coping with mediocre propagation conditions.

The Page Mode

The PAGCH configuration may change in time to adapt to the traffic distribution. Although this configuration is managed by the BSC, it is usually the OSS which decides on such a change, either automatically on the basis of traffic observations or through the command of an operator. The configuration of the PAGCH in a given cell must be known by all mobile stations camped on this cell, and is therefore part of the broadcast information. When the configurations changes, the BSC must co-ordinate the change of the broadcast information and the scheduling of paging messages. It is not possible to control very precisely the time at which each mobile station will decode the new broadcast parameter, and hence switch to the new configuration. The corresponding transitional period may last up to a few seconds. In order to avoid the loss of paging messages during this "fuzzy" period, a special feature has been incorporated in the *Specifications*: the page mode. The page mode indicates to mobile stations in exactly which part of the PAGCH their own paging messages may be sent.

The three values of the page mode are the following:

- the "normal" page mode corresponds to the basic scheme. Paging messages are sent only on the sub-channel as defined by the PAGCH configuration and the IMSI;
- the "full" page mode has been designed to cope with a dynamic change in the PAGCH configuration. When this mode is indicated to mobile stations of a sub-group, then it means that paging messages for the subscribers of this sub-group may be sent anywhere on the PAGCH of the same *timeslot*;
- The "next-but-one" page mode has been introduced for sophisticated scheduling algorithms. It enables the BSS to send additional paging messages for subscribers in a given sub-group in another paging sub-channel. This feature may be useful in a situation of temporary overload on some of the sub-channels, or to free a block to send an initial assignment message. Based on the fact that the definition of paging sub-channels includes an implicit numbering of these sub-channels, the "next-but-one" mode indicates that paging messages for mobile stations

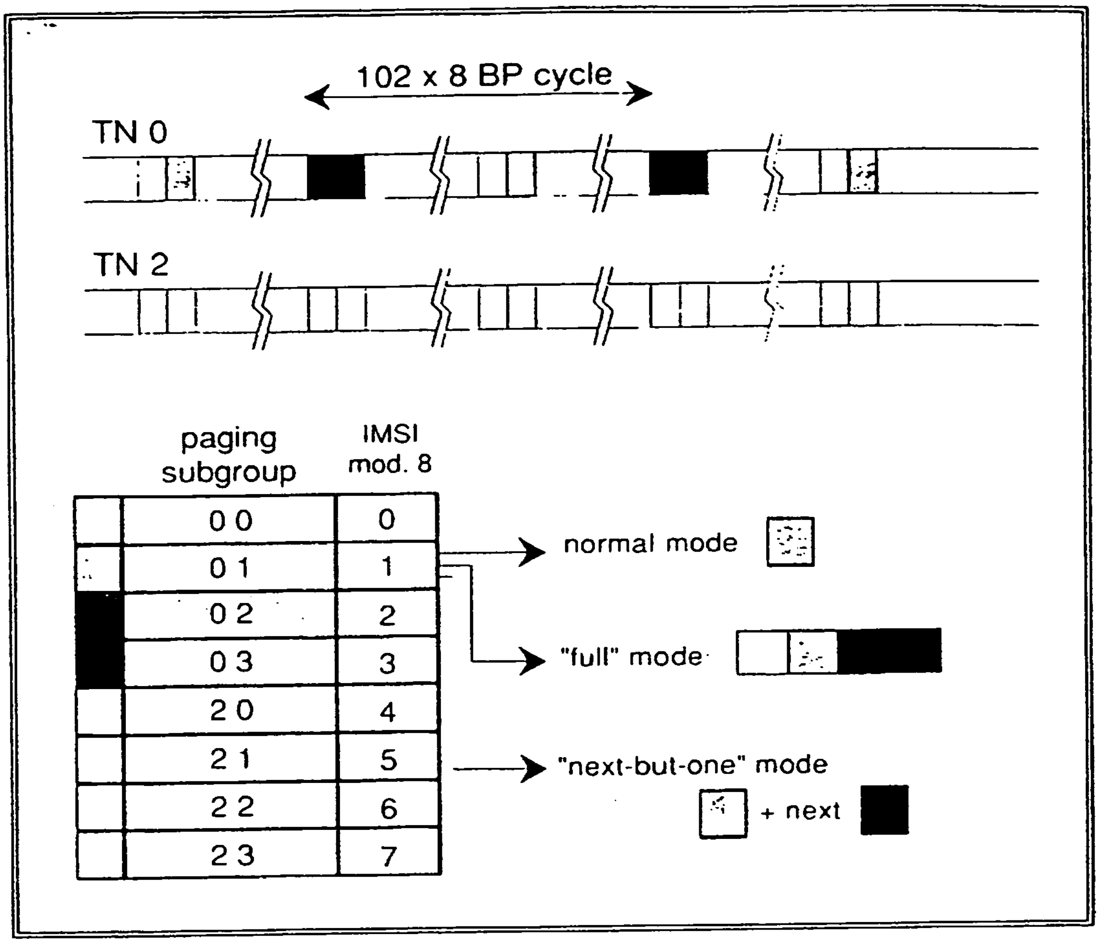


Figure 6.23 – Different page modes

In addition to the "normal" paging mode, the "full" page mode and the "next-but-one" page mode are useful in transition situations, or to compensate for partial congestion situations.

normally paged in sub-channel *n* (the sub-channel where the page mode indication was sent) will in addition be found in the next block of the paging sub-channel *n*+2, modulo the total number of paging sub-channels on the current *timeslot*. The example of figure 6.23 will help to understand this feature.

It is worth noting that the mobile station is free to listen to more than the required minimum. This may indeed represent a simplification for vehicle-mounted mobile stations, for which battery consumption is not at stake. In cases where the load on the PAGCH is very low, it might even be interesting for the BSS to send the paging messages in more blocks than those of the required paging sub-channel. This will allow a speed-up of their reception for mobile stations which listen to more than just their own paging sub-channel.

6.3.3. PROCEDURES FOR TRANSMISSION MODE AND CIPHER MODE MANAGEMENT

Transmission modes, as far as radio resource management is concerned, involve 7 different types in phase 1 and are expected to expand to 12 in phase 2 (see table 6.1, page 321). The mode is one of the properties of the transmission chain. Other variable characteristics of this chain include the cipher mode (whether transmission is ciphered or in clear text), as well as both uplink and discontinuous transmission modes.

At initial assignment, the transmission mode is chosen by the BSC. It consists necessarily of one of the "signalling only" modes, in clear text. The channel may be a TACH/8 or a TACH/F. Afterwards during the lifetime of the RR-session, the choice of the transmission mode depends on the communication needs and is done by the MSC, which can request a change of the transmission mode at any time while a connection is established. It does so through an "assignment" procedure (which does not necessarily result in a radio channel assignment). The basic procedure consists of a BSSMAP ASSIGNMENT REQUEST message, describing the transmission characteristics as desired by the MSC, and the corresponding BSSMAP ASSIGNMENT COMPLETE acknowledgement, as shown in figure 6.25. A negative answer is also possible in case of problems, and is carried in a BSSMAP ASSIGNMENT FAILURE message.

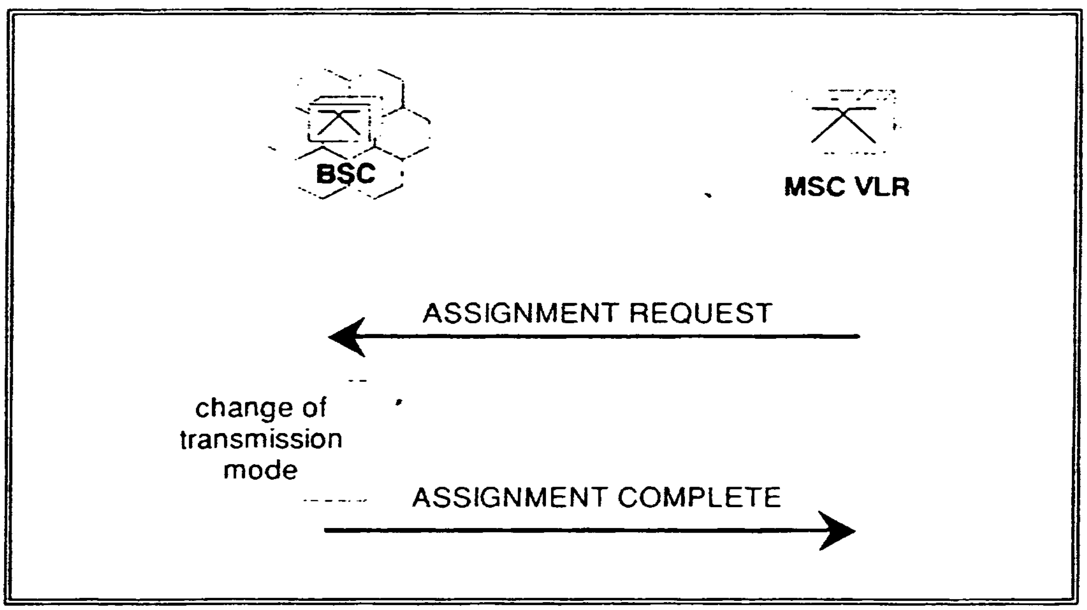


Figure 6.24 – Change of the transmission mode by the MSC

The MSC may change the transmission mode of an RR-connection at any time, by running the "assignment" procedure towards the BSC, which then takes charge of controlling the transmission mode change.

Another special case is when the allocation cannot proceed immediately, but the request is put in a queue; the MSC can be warned of the situation by a BSSMAP QUEUING INDICATION message (the completion or failure indication is eventually sent later). The request also includes the identity of the newly allocated terrestrial circuit between the BSC and the MSC, when required, i.e., when the mode is changed from "signalling only" to another one.

The action of the BSC when receiving a BSSMAP ASSIGNMENT REQUEST message depends on the comparison between the existing transmission mode and the required one:

- if both modes are the same, the BSC just sends back the BSSMAP ASSIGNMENT COMPLETE message back to the MSC without any other action;
- if both modes differ by the type of information to be transmitted, but use the same type of channel, the BSC performs a "mode modify" procedure before acknowledging the MSC request;
- if the new mode requires a channel of a type different from the one in use, the BSC performs a subsequent assignment procedure, i.e., it transfers the connection to a channel of the required type before acknowledging the MSC request.

The two last cases will now be studied in more detail.

6.3.3.1. The Mode Modify Procedure

As for any procedure affecting the transmission mode, the mode modify procedure includes two parts: the configuration of the transmission devices on the infrastructure side (BTS, TRAU and BSC), and the configuration of the mobile station. No means are provided to synchronise these two parts in a precise way, resulting usually in a short period of time during which the whole configuration is inconsistent. The *Specifications* do not specify in which order the two configuration steps should be managed, and this choice may have an impact on the period of inconsistency. In the following example, it has been hypothesised that the two tasks are run in parallel.

The BSC triggers the reconfiguration of the BTS and the TRAU by sending an RSM MODE MODIFY REQUEST to the BTS. Following reception of this message, the BTS modifies its coding and decoding algorithms and changes the in-band mode information in the BTS-TRAU frames. The TRAU, in turn, reacts by modifying its data processing (speech

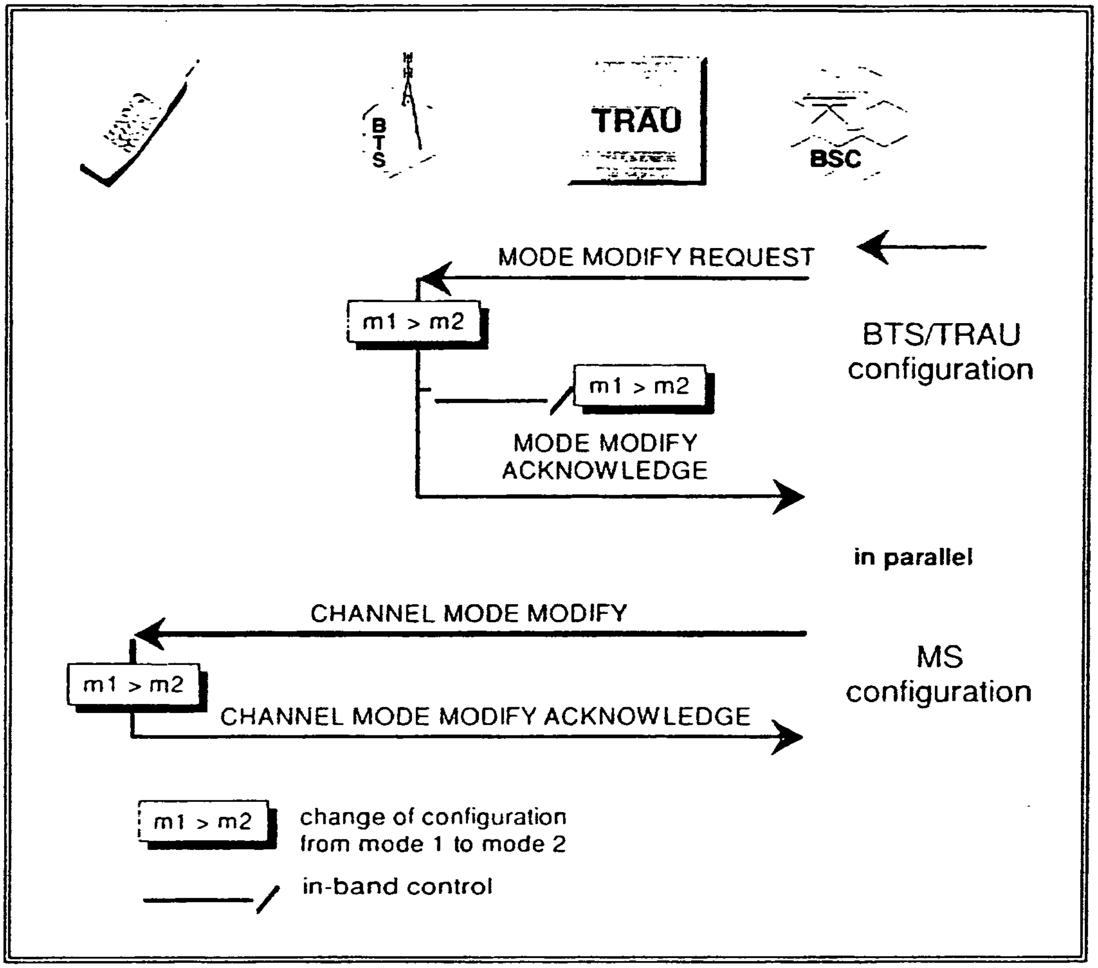


Figure 6.25 – The "mode modify" procedure

The BSC is in charge of configuring both the BTS and the mobile station, but the order in which these steps should be performed is not specified. The TRAU is configured through in-band information from the BTS.

coding or data rate adaptation). If the new mode is speech, then synchronisation between TRAU and BTS is needed. When the chain is ready (the *Specifications* do not specify whether this includes the synchronisation with the TRAU or not), the BTS answers the BSC by sending back an RSM MODE MODIFY ACKNOWLEDGE message.

In parallel, the BSC triggers the reconfiguration of the mobile station by sending an RIL3-RR CHANNEL MODE MODIFY message containing the new mode to be applied (see also figure 6.25). When the mobile station receives the order, it modifies its channel/source coding/decoding according to the new requirements and answers with an

RIL3-RR CHANNEL MODE MODIFY ACKNOWLEDGE message to the BSC via the BTS. It is worth noting that the connection of higher level devices (microphone and loudspeaker for speech, or terminal adapter for data) is not controlled by this procedure, but by a call control procedure (see Chapter 8).

A third action needs to be performed by the BSC: circuit switching. It may happen that nothing needs to be done, for instance when a terrestrial circuit is already established and satisfies the needs. But if the BSC-MSC terrestrial circuit needs to be reset or altered, the BSC must connect it to the correct BTS-BSC circuit (in correspondence with the TACH/F). The most complex case arises when the TRAU is remote, and must be changed to another one, adapted to the new mode. This is in fact only possible when BSC-controlled switching facilities exist between both the BTS and TRAU and between the TRAU and BSC. In this case, the BSC must switch the connections before ordering the reconfiguration of the BTS. In the other cases, BSC switching can be done in parallel with the other actions.

6.3.3.2. The Subsequent Assignment Procedure

When a change of the radio channel is required in addition to the above-described "mode modify" procedure, the procedure is somewhat more complex, because the change of channel implies a break in the signalling carrying capability between mobile station and infrastructure. Besides, let us recall that the BTS devices in charge of transmission on a given TACH are independent and do not communicate. Therefore, the whole control of the operation is centralised in the BSC, and the operation itself is quite similar to a handover.

A transfer of channel first starts with the setting up of the new path in the infrastructure. This includes the allocation of a new radio channel (with all the priority and queuing management aspects described on page 357, to cope with congestion), the activation of the corresponding BTS device, and possibly the allocation of a TRAU and the switching necessary to connect all these terrestrial segments.

The activation of the BTS is ordered by the BSC through a simple request/acknowledgement procedure, as shown in figure 6.26. The RSM CHANNEL ACTIVATION message contains all the information specifying the transmission mode, including the basic transmission mode (among those listed in table 6.1), the cipher mode and the downlink discontinuous transmission mode. An uplink discontinuous transmission mode is also

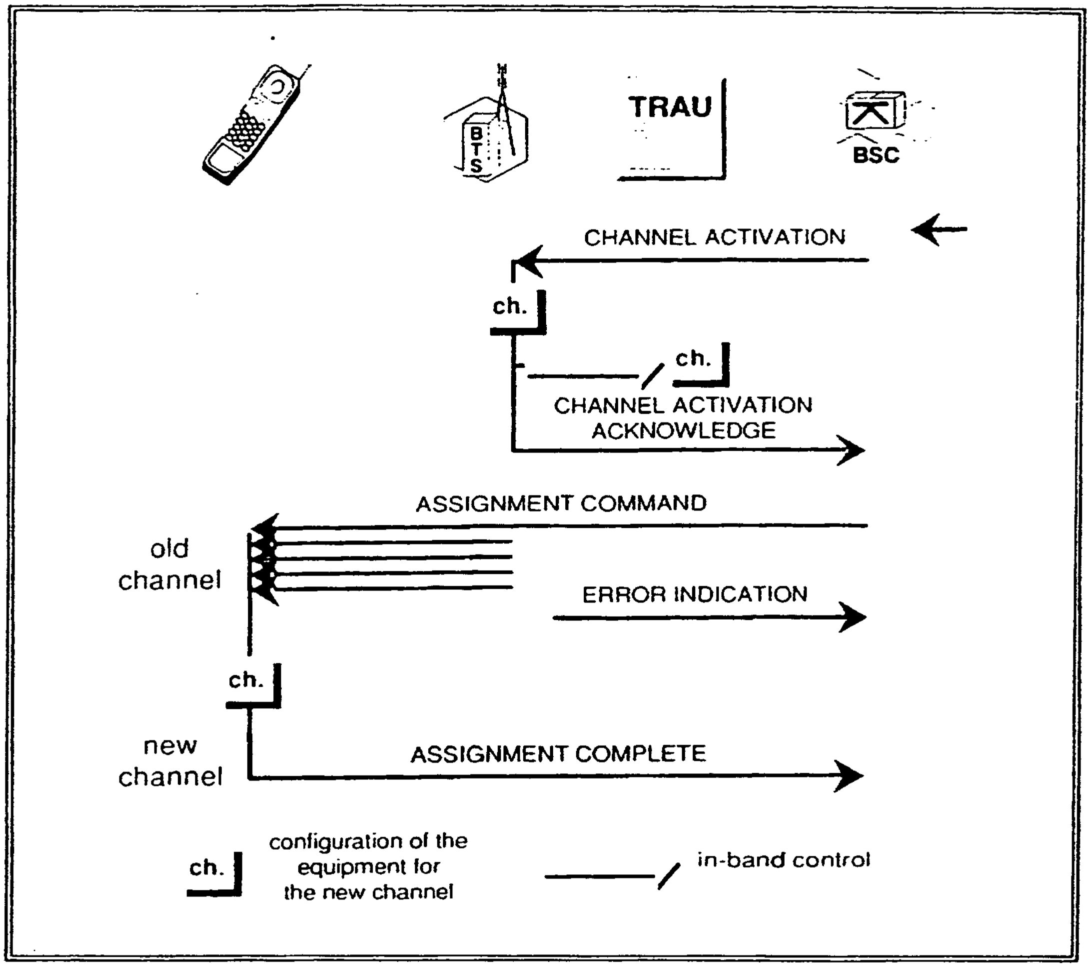


Figure 6.26 – Activation of a new channel in the BTS

After the activation handshake on the Abis interface, the BSC orders the mobile station to change channel by an RIL3-RR ASSIGNMENT COMMAND message, which is in general not acknowledged by the mobile station on the old channel. The completion of the procedure is done on the new channel after establishment of the full signalling link.

sent, though one may wonder what it may be used for, and moreover what its meaning is, since in some cases the network leaves the choice to the mobile station. In addition, it contains the information needed by the mobile station for access (see the initialisation of the timing advance, page 346) and the first power control settings. The BTS, upon reception of this message, starts the in-band information exchanges with the TRAU, to set the basic transmission mode and the discontinuous transmission modes; this is the point where synchronisation with the TRAU starts.

Once the BTS and the TRAU are activated (more exactly when the BSC has received the RSM CHANNEL ACTIVATION ACKNOWLEDGE

message), the BSC orders the mobile station to perform the transfer of channel, through an RIL3-RR ASSIGNMENT COMMAND message. The previous path, including the signalling connection, is not released by the infrastructure at this moment. This allows the mobile station to go back to the previous channel should access fail on the new channel for any reason. In the case of a timed assignment (see page 353), the mobile station stays on the old channel until the instant of change indicated by the infrastructure. Otherwise, the mobile station performs the transfer immediately after reception of the RIL3-RR ASSIGNMENT COMMAND message, not even acknowledging the corresponding frame at layer 2. This lack of acknowledgement results in a repetition of the message by the BTS, until it decides that a link layer failure has happened, of which the BSC is advised. The BSC does not act on this indication, since it knows that an assignment is in progress. In case of return to the old channel, the mobile station starts re-transmitting on this channel by performing a link establishment, which resets all contexts irrespective of what happened in the link layer process in the BTS. The transmission mode used after return on the old channel is the one that was used on it, not the one asked for in the assignment message. Whether in the case of a successful procedure or of return to the old channel, the interruption of the link layer may result in leaving a message sent by the mobile station in a non-acknowledged state. This situation is handled by the upper layers, as will be explained below.

Once the mobile station has changed its settings to the ones corresponding to the new channel, it starts transmission and reception according to the transmission mode indicated in the RIL3-RR ASSIGNMENT COMMAND message. It must also establish a new signalling link in acknowledged mode. The *Specifications* do not say whether this link layer establishment should be done first (before the transmission of any user data) or not. Anyway, once this link is established, the mobile station sends an RIL3-RR ASSIGNMENT COMPLETE message to the BSC before any other message. Then, all the messages waiting for transmission can be sent: first the ones already sent on the previous channel but still not acknowledged, then the ones which have arisen during the procedure. The same applies in case of return to the old channel, except that the first message is an RIL3-RR ASSIGNMENT FAILURE.

A message sent by the mobile station before link interruption and not acknowledged then cannot be lost, but may be duplicated. RIL3-RR messages are considered immune to duplication, and therefore no mechanism has been introduced to cope with such duplications. This is worth remembering when new procedures are added in the future! The

way in which the BSC is implemented has also some impact. A "careful" BSC will postpone the sending of an RIL3-RR ASSIGNMENT COMMAND (or an RIL3-RR HANDOVER COMMAND) while it is waiting for the answer to another RR command (e.g., for the start of ciphering!).

In the case of upper layers, e.g., call control or mobility management, message duplication could in some cases be harmful. In order to avoid problems, correction is obtained by detection and suppression of the duplicated message. Detection is obtained through the use of a simple numbering scheme, the 1-bit sequence number referred to as N(SD) in the *Specifications*. This number is included in each upper layer message, and is changed for each new message. Since messages are sent and acknowledged one at a time (window size 1), this scheme is enough for the infrastructure to detect duplication. When it receives two successive messages with the same N(SD), it discards the second one. This task is performed at the RR level, but not in the BSC. It is one of the tasks of the anchor MSC, which is the destination of upper layer messages sent by the mobile station. The reason why it cannot be performed in the BSC is that it must be done at a point which remains stable when the transmission chain changes, and the reader will recall that the only stable part of an RR-session lies in the anchor MSC.

6.3.3.3. The Change of the Cipher Mode

During the lifetime of an RR-session, the cipher mode may change on the radio interface. This procedure is not all that easy. As explained in Chapter 4, the cipher mode is applied to all transmitted information, including signalling messages. Thus, a change of the cipher mode entails a signalling break, with a possibility of message loss.

In the case of subsequent assignment and handover, the problem raised by the signalling break is solved through a full re-establishment of the signalling link. This is full-proof against call loss only because both the old and the new channels are available during the critical period, thus allowing the return to the old channel in case of problems on the new one. In the case of a cipher mode change, it would be too costly to require the BTS to perform reception in both modes (ciphered and non-ciphered) simultaneously, and a different solution was adopted. It consists of dividing the procedure into three steps instead of two:

- step 1: the BTS is configured to transmit according to the old mode, and receive according to the new mode;
- step 2: the mobile station is configured to the full (transmission

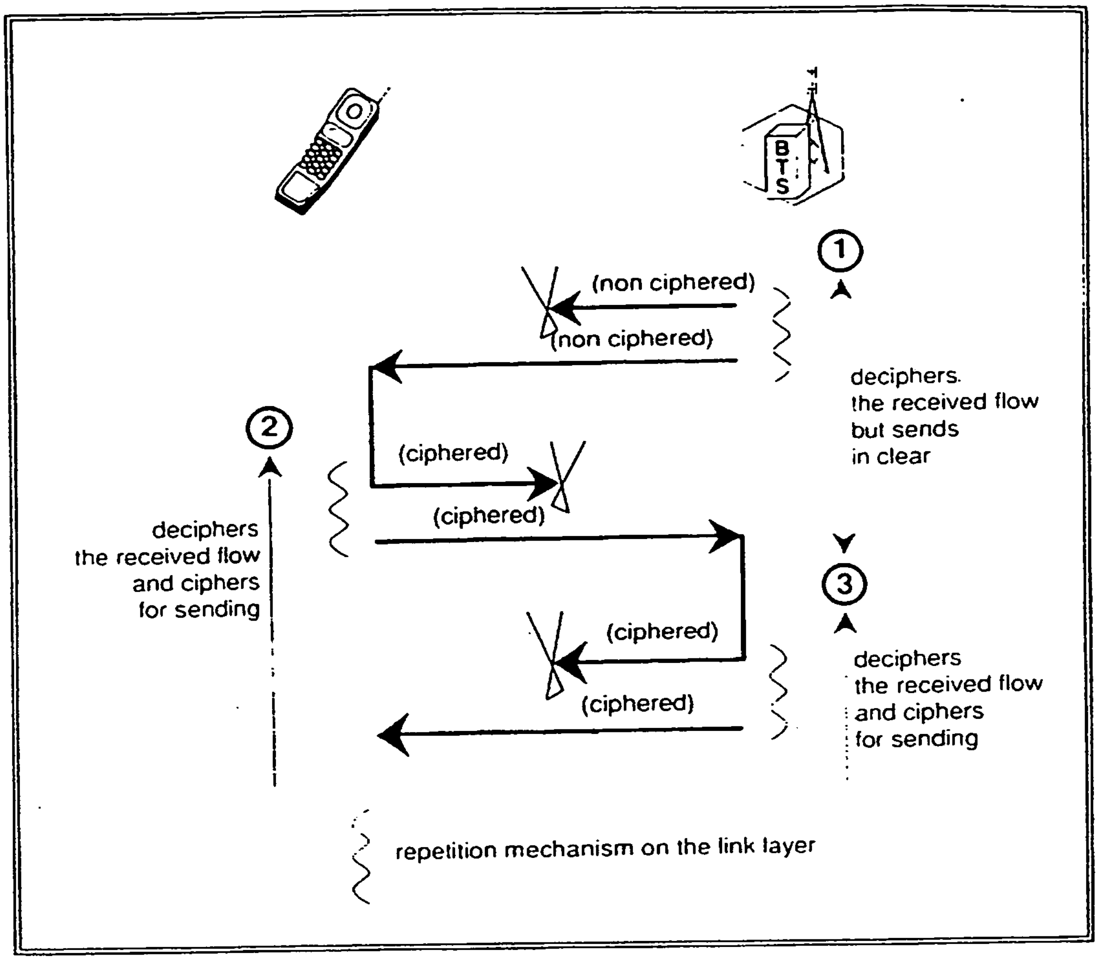


Figure 6.27 – Cipher mode change: the 3 steps

In order to cope with an interruption in the signalling link when a change of the cipher mode occurs, the procedure is split into three steps, so that only one direction of the transmission will be in a critical state at each transition. The example shown here represents a transition from clear text to ciphered mode.

and reception) new mode;

- step 3: the BTS is configured to the full new mode.

It can be accepted that steps 1 and 2 be inverted, but this results in an increased frame loss probability, since in this case both uplink and downlink messages would be lost between step 1 and step 2, whereas only downlink messages are lost in that phase with the given order.

The critical period is split in two with such a mechanism. From the first to the second step, BTS to MS transmission functions correctly, but not in the other direction. This is sufficient for a downlink message

triggering step 2 to be repeated a number of times by the infrastructure, until received. Similarly, from step 2 to step 3, only MS to BTS transmission is correct. But this is sufficient for the mobile station to repeat the uplink acknowledgement message required after step 2, to trigger step 3. In no case does a single message loss jeopardise the whole connection.

The procedure, shown in figure 6.27, relies heavily on the link layer mechanisms, namely the repetition of messages after a given timeout period, in the absence of an acknowledgement from the other side. Though not obvious, it can be shown through a careful study that, even with a window size greater than 1, and/or messages waiting to be sent on the mobile station side, the procedure terminates correctly even if the radio conditions are not perfect. Yet it should be noted at this point that the cipher mode change procedure can by itself cause the loss of a frame, and transmission is therefore marginally more sensitive to errors at the time of a cipher mode change.

Because of the strong requirement on the order of the three steps listed above, and because ciphering is implemented by the BTS on the infrastructure side, the procedure is not managed by the BSC, but by the BTS. The BSC transmits a single order to the BTS, which then runs the procedure, including its own configuration as well as the one for the mobile station. As a matter of fact, the decision to change the cipher mode is taken by the MSC, and results in a cascade of messages from MSC to BSC, then from BSC to BTS, and lastly from BTS to MS on the radio interface. The whole chain is shown in figure 6.28.

The BSSMAP CIPHER MODE COMMAND message indicates the new requested mode. After having extracted the new parameters from this message, the BSC builds up an RIL3-RR CIPHERING MODE COMMAND message targeted at the mobile station and encapsulates it in an RSM ENCRYPTION COMMAND message sent to the BTS.

The BTS then configures its reception to the new mode and sends the encapsulated RIL3-RR CIPHERING MODE COMMAND message to the mobile station using the old mode. When receiving it, the mobile station sets its configuration to the new mode and puts an RIL3-RR CIPHERING MODE COMPLETE message in the sending queue. This message is not necessarily the first one to be sent in the new mode, as another one may be there, and the RIL3-RR CIPHERING MODE COMPLETE message has no particular priority. The main reason for the existence of this acknowledging message is to ensure that at least one layer 3 message is sent at this moment, in order to trigger the switching to the new mode in

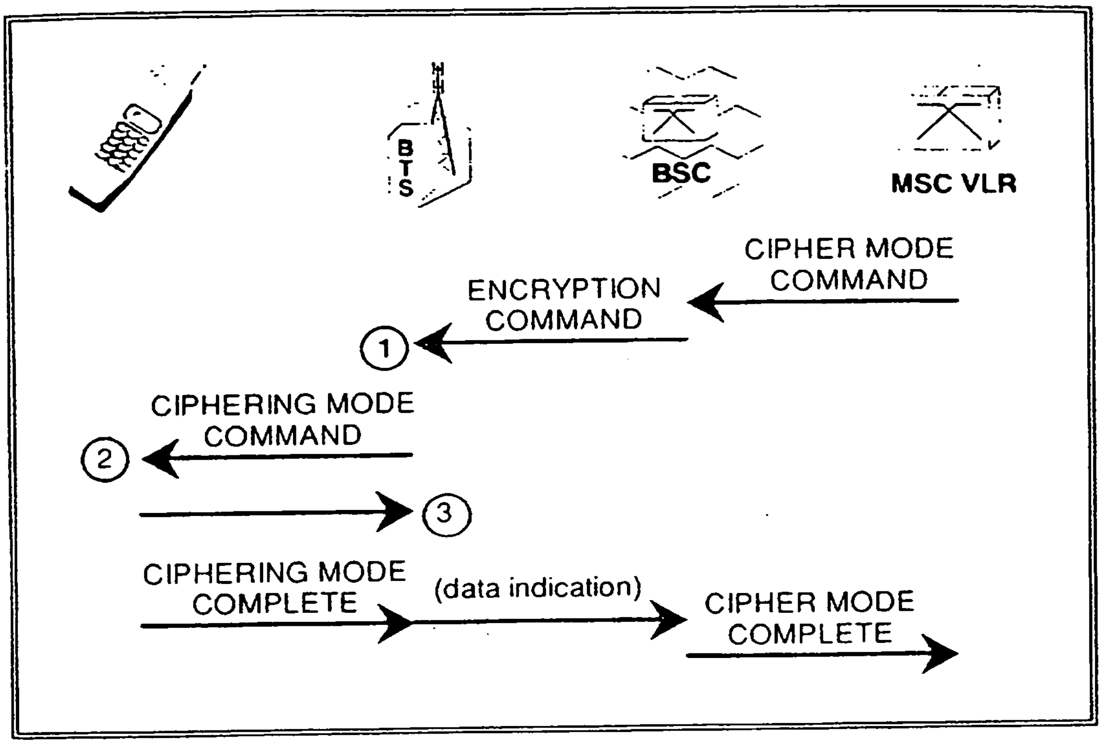


Figure 6.28 – The cipher mode setting procedure

The procedure is initiated by the MSC, but all the synchronisation management is done by the BTS, which is in charge of ciphering/deciphering. The grey area corresponds to this management, as detailed in figure 6.27.

the BTS. It is when receiving any correctly decoded message (in the new mode), which implies that the mobile station has indeed correctly switched to the new mode, that the BTS switches fully to the new mode. Either then or afterwards, the RIL3-RR CIPHERING MODE COMPLETE message is forwarded to the BSC, which translates it into a BSSMAP CIPHER MODE COMPLETE message to indicate to the MSC that its request has been fulfilled.

The case when the MSC requires a new mode which is already the one in place has not been dealt with. This omission is purposeful, since in this case the *Specifications* impose that the whole procedure be run anyway. This requirement comes from the second meaning of the RIL3-RR CIPHERING MODE COMMAND message, which is used to acknowledge an RIL3-MM CM SERVICE REQUEST (this will be explained in Chapter 7). The cipher mode command procedure is then strictly speaking not only spread over the RR layer and the link layer, but also on the MM layer.

It is not possible to change the cipher mode when changing the channel (assignment or handover) for unclear reasons (later phases may indeed remove this constraint). A mobile station developed according to the phase 1 *Specifications* must apply the same cipher mode on the new channel as was used on the previous one, and this situation can lead to problems in case of collision between a channel transfer and a change of the cipher mode. In order to avoid such problems, a sequential approach between these two procedures must be sought by the BSC.

6.3.3.4. Discontinuous Transmission Modes

The *Specifications* do not indicate clearly whether it is allowed to change the discontinuous transmission modes during the lifetime of an RR-connection. The need can be identified, at least for the downlink discontinuous transmission mode, since this mode may depend on the correspondent and a single RR-session can be used for several communications in succession. It should be noted as a preliminary remark that there is no need for the receiver to know beforehand whether the sender applies discontinuous transmission or not. Therefore, no specific procedure is required for the indication of the downlink discontinuous transmission mode to the mobile station or of the uplink discontinuous transmission mode to the infrastructure transmission devices. Means exist for the other cases, and will now be described.

As far as the downlink discontinuous transmission mode is concerned, it must be ordered connection by connection by the MSC to the BSC, which configures the BTS, itself configuring the TRAU. It must be recalled at this stage that, on the transmitting side, discontinuous transmission does not only result in some frames not being sent, but it also modifies the speech coding algorithm (e.g., sending of comfort noise frames is done only in discontinuous transmission mode). The initial command is issued through a downlink discontinuous transmission indicator included in the messages used for the management of the basic transmission mode from the MSC (BSSMAP ASSIGNMENT REQUEST) and for channel activation toward the BTS (RSM CHANNEL ACTIVATION). The TRAU, in turn, is configured through an in-band indicator set by the BTS. Means to change the downlink discontinuous transmission mode on an established RR-connection also exist, since the same indicator is also included in the RSM MODE MODIFY REQUEST. Besides, the BSSMAP ASSIGNMENT REQUEST message can be used by the MSC to trigger a discontinuous transmission mode change on the infrastructure side. As already mentioned, the mobile station need not be warned of such a

change, and no procedure has been defined on the radio interface for this purpose.

The mobile station can be ordered to use the discontinuous transmission mode in the uplink direction as a cell option. This view is consistent with the consideration of discontinuous transmission as a means to improve spectral efficiency, but it does not take into account mobile stations or connections at an individual level. The cell options are regularly broadcast on the BCCH for mobile stations in idle mode, and they are also part of the general information sent to mobile stations on their SACCH when they are in dedicated mode. The uplink discontinuous transmission mode could thus at first sight be set on a connection basis by this slow signalling method, and changed in the same way. However, the specification of the RSM protocol requires to set the general information messages sent on the SACCH on a transmitter/receiver (TRX) basis. This state of things is even more fuzzy because of the presence of an uplink discontinuous transmission indicator (besides the downlink one) in the *CHANNEL MODE* information element, which is contained in the messages for the management of the basic transmission mode between BSC and BTS. The BTS has no need to know whether uplink discontinuous transmission is used or not; a possible justification for the presence of this indicator would be to ask the BTS to modify the RIL3-RR SYSTEM INFORMATION 6 it sends to the concerned mobile station. A point against this interpretation is that the information in the SYSTEM INFORMATION TYPE 6 message has three values (DTX must be applied, must not be applied or may be applied), whereas the information in the *CHANNEL MODE* information element is two-valued.

6.3.4. HANDOVER EXECUTION

The handover execution procedure enables the network to command a mobile station in dedicated mode to go onto another channel in another cell. The handover execution procedure is very close to the subsequent assignment procedure. The fundamental difference is the change of cell. The handover execution procedure differs mainly from the subsequent assignment procedure by the timing advance management, by the need to transmit some data specific to the new cell, and by a few limitations.

Basically the procedure was designed not to constrain the type or the mode of the new channel. As for the assignment procedure, the type and mode of channel before the handover, or the type and mode of the channel after the handover can be anything. Or almost... If the indicative value of timer T3124 (a timer used on the mobile station side for handover between non-synchronised cells) as defined in TS GSM 04.08 is applied by mobile stations in all cases, handover toward a TACH/8 does not run properly; it is therefore not part of the test requirements for type approval of phase 1 mobile stations. This will be corrected in phase 2. Yet a handover from a TACH/8 toward a TACH/F is possible. This has

some application. When a connection is established, the cell has been chosen by the mobile station. While in connected mode, the cell is chosen by the network. The two cell choice algorithms are different, and hence may lead to different results in many cases. When this happens, and the initial channel is a TACH/8 whereas the needed channel is a TACH/F, a handover directly from the TACH/8 to a TACH/F in the right cell is quicker than performing a subsequent assignment first, and a handover next. This way of proceeding is called a "directed retry" (from the name of a similar action in analogue networks), and can be used also for congestion cases.

The handover procedure can be executed for different reasons, explained within the scope of the handover preparation function. But in all cases, the decision to attempt the handover of a given mobile station is taken by the BSC. Once so done, and once the new cell (or a list of candidate new cells) has been chosen, the actual transfer must be co-ordinated between the mobile station and the machines managing the old cell (BTS-old) and the new cell (BTS-new).

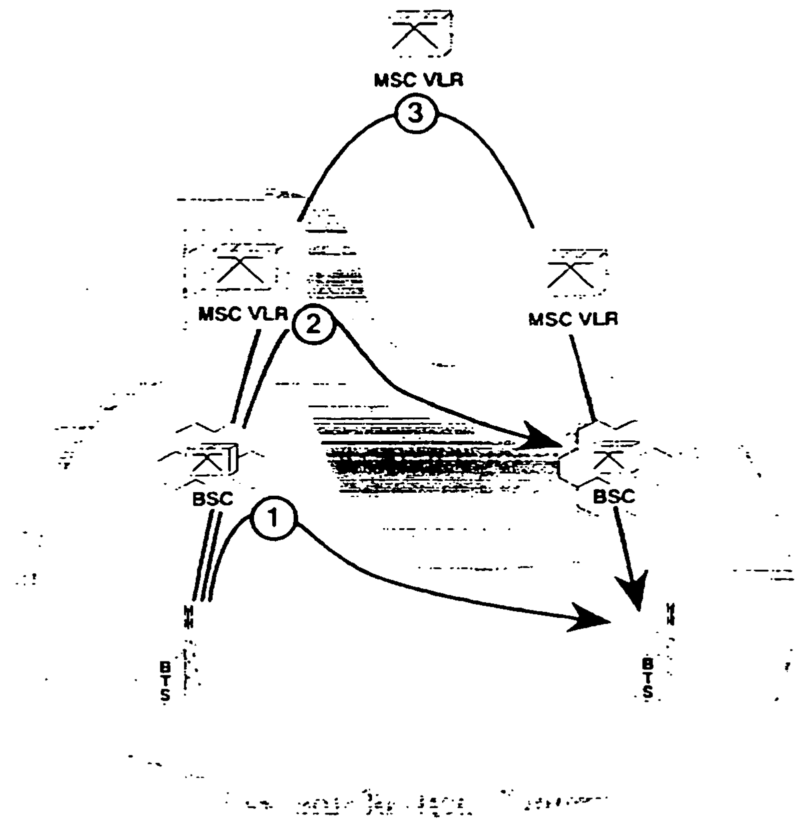
The handover procedure comes in several varieties, according to two main criteria.

The first criterion is related to the timing advance issue, and impacts only the "incoming" part of the radio interface procedure, between the mobile station and BTS-new. As expressed in the functional description of the timing advance management (see page 346), two cases can be distinguished:

- the mobile station is able to compute the new timing advance (to be applied with BTS-new), because the old and the new cells are synchronised (**synchronous** handover);
- the timing advance must be initialised both at the mobile station and at BTS-new during the handover procedure (**asynchronous** handover).

In a way, the subsequent assignment procedure can be considered as a third case, in which the timing advance is not changed.

The second criterion concerns the location of the switching point in the infrastructure. This location impacts heavily the procedures to be used between infrastructure entities; the procedure on the radio path is not impacted by such a consideration, except to distinguish the special case of intra-cell handover which uses the same procedure as for subsequent channel assignment.



1 **BSC** : intra-BSC handover

2 **MSC** : inter-BSS, intra-MSC handover

3 **anchor MSC** : inter-MSC handover (when MSC-old = anchor MSC), or subsequent inter-MSC handover (otherwise).

Figure 6.29 – Position of the switching point at handover

The switching point depends on the relative position in the infrastructure hierarchy of the BSCs controlling the two cells. In all cases, the anchor MSC remains involved in the communication path, and there can be but only one other MSC in the path.

In order to describe the different cases, the suffix "old" shall be used to refer to all functional entities along the communication path before the handover, and "new" shall be used for the path after handover. BTS-old, BSC-old and (transit) MSC-old represent the machines in charge of the old cell, and BTS-new, BSC-new and (transit) MSC-new the machines in charge of the new cell. It may happen that BSC-old = BSC-new, or MSC-old = MSC-new. Regarding the relay MSC, it

must be recalled that the anchor MSC, which originally established the RR-session, may be the same as MSC-old or MSC-new, or both, but may also be different. If the call already goes through two MSCs, the anchor MSC and MSC-old are distinct. If a handover is needed toward a cell of a new MSC different from MSC-old, the switching point is the anchor MSC. Figure 6.29 shows the most complex case, when all entities are physically distinct.

Whether synchronous or asynchronous, whether inter- or intra-MSC, and whether inter- or intra-BSC, the execution of handover is composed of two main phases:

- in a first phase, BSC-old triggers a set of events with the purpose of establishing the future communication path. Once this is done, this phase terminates with the sending of a handover command to the mobile station;
- in a second phase, the mobile station accesses the new channel. This access triggers the switch of paths in the infrastructure, and the release of the old path.

6.3.4.1. The Set-Up of the New Path

Once the decision of handing over a given communication has been taken by BSC-old, this must be indicated to the switching point. The latter must in turn establish the terrestrial resources, if need be, up to BSC-new, signal to it to allocate a radio resource and more generally provide all impacted machines with the information they need for the handover and the future management of the connection. This information includes:

- the transmission mode, used to choose and configure the transmission path in an appropriate way, including the new radio channel;
- the cipher mode;
- the identity of the origin cell, used to determine whether the handover can be done in a synchronous or an asynchronous way;
- the mobile station classmark, used for future management of the connection.

Once BSC-new is aware of all this information, it is in a position to allocate the new channel, to build the RIL3-RR HANDOVER COMMAND message and to transmit it to the switching point, which in turn will

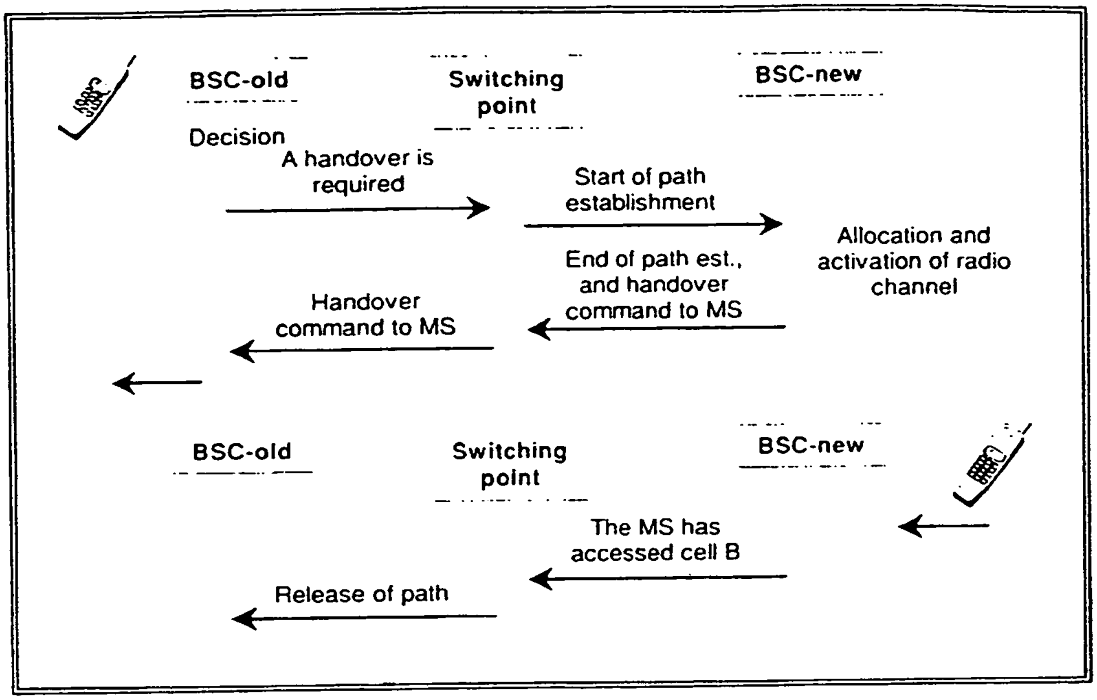


Figure 6.30 – Handover execution sequence

Once the handover decision taken, the set-up of the new path unfolds in four steps, to prepare for mobile station access.

convey it to the mobile station along the old path, as shown in figure 6.30. Let us study these steps and the respective variations in detail.

From BSC-old to the Switching Point

The purpose of this exchange is the transmission of the information that a handover is needed, and toward which cell (or cells). The different cases depend on the nature of the switching point (see figure 6.31):

a. BSC-old is the switching point (BSC-old = BSC-new):

   this step is internal and does not raise any problem.

b. MSC-old is the switching point (BSC-old ≠ BSC-new, MSC-old = MSC-new):

   in all cases when the target cell is not under its control, BSC-old sends a BSSMAP HANDOVER REQUIRED message to MSC-old, containing the identities of the target cell(s) and of the origin cell.

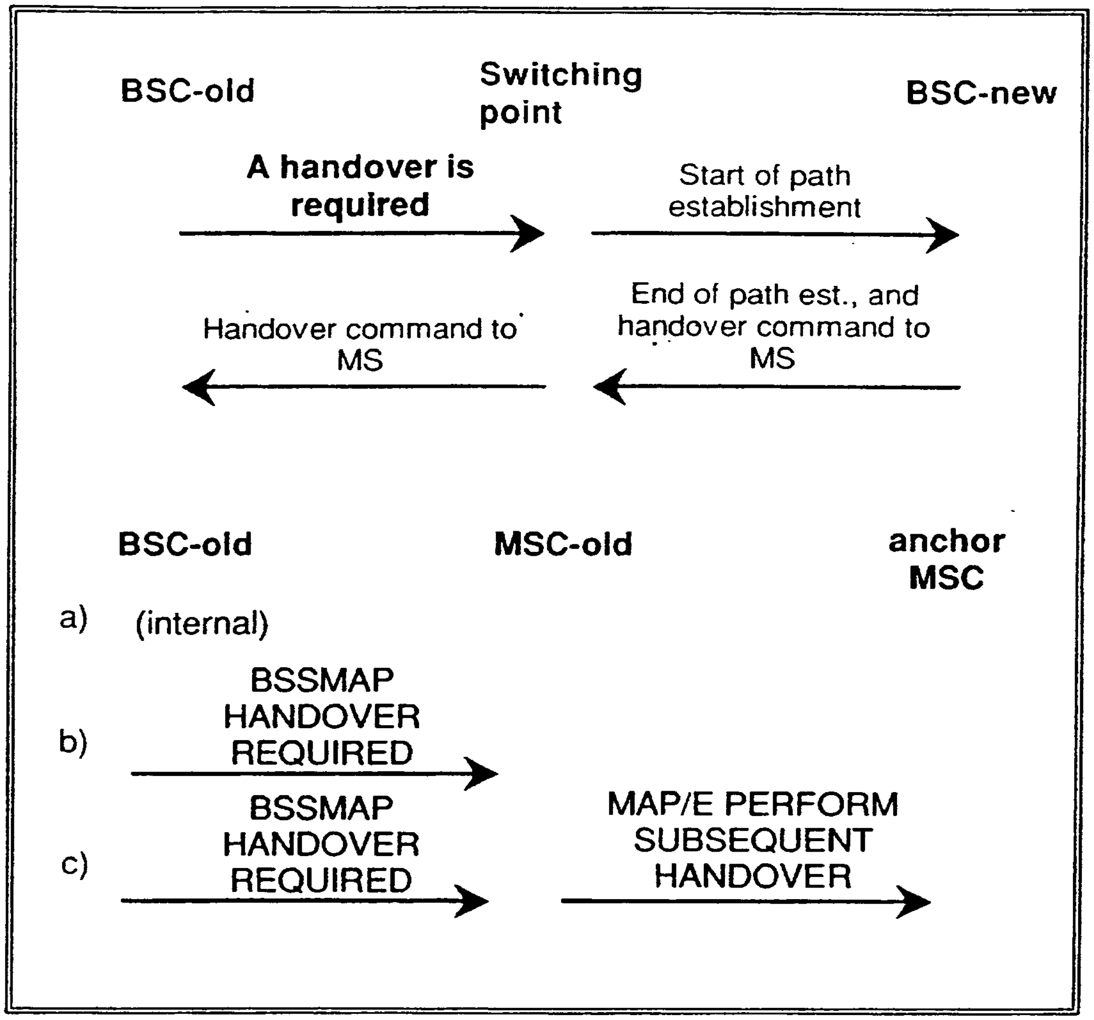


Figure 6.31 – Handover requirement from the serving BSC to the switching point

When applicable (depending on the relative position of the switching point), the BSSMAP HANDOVER REQUIRED message may be followed by a MAP/E PERFORM SUBSEQUENT HANDOVER message.

c. Anchor MSC is the switching point, and differs from MSC-old:

the behaviour of BSC-old is the same as in the previous case, but the behaviour of MSC-old when receiving the BSSMAP HANDOVER REQUIRED message is different. It translates the message in a MAP/E PERFORM SUBSEQUENT HANDOVER message towards the anchor MSC. Both messages have similar contents.

From the Switching Point to BSC-new

The purpose of this step is to establish the signalling pathway between the switching point and BSC-new, to prepare for the establishment of the circuit if need be and to provide the machines along the new path with the information they need. The events at this stage depend on the relative position of the switching point and BSC-new, as shown in figure 6.32:

a. BSC-new is the switching point (BSC-old = BSC-new):

   this step is internal. The BSC is aware of all the relevant information since it already manages the current context of the connection. No terrestrial circuit needs to be allocated apart from the Abis circuit linked to the new radio channel.

b. MSC-new is the switching point (BSC-old ≠ BSC-new, MSC-old = MSC-new):

   when receiving the indication that a handover is required, MSC-new establishes an SCCP connection towards BSC-new. This connection will be used throughout the life of the new RR-connection. MSC-new then transmits a BSSMAP HANDOVER REQUEST message to BSC-new, including the information on the cells (both origin and target cells), the transmission mode (derived from the present needs, hence possibly differing from the characteristics of the connection established through the old cell), the cipher mode (which *must* be the same as in the old cell, since the mobile station will assume so), the classmark and finally the reference of the terrestrial channel between MSC-new and BSC-new if need be.

c. Anchor MSC is the switching point, and differs from MSC-new:

   this corresponds to the most complex case. The tasks of the anchor MSC cannot be done in a single step as in the previous case, because the new communication path between the anchor MSC and BSC-new (through MSC-new) may transit via the PSTN or the ISDN. Standard inter-switch procedures must therefore be used, which are part of TUP or ISUP (or national variants of them). These protocols do not provide the means to convey the relevant GSM information. They are therefore only used for circuit set-up and MAP/E procedures are used for the specific handover signalling needs.

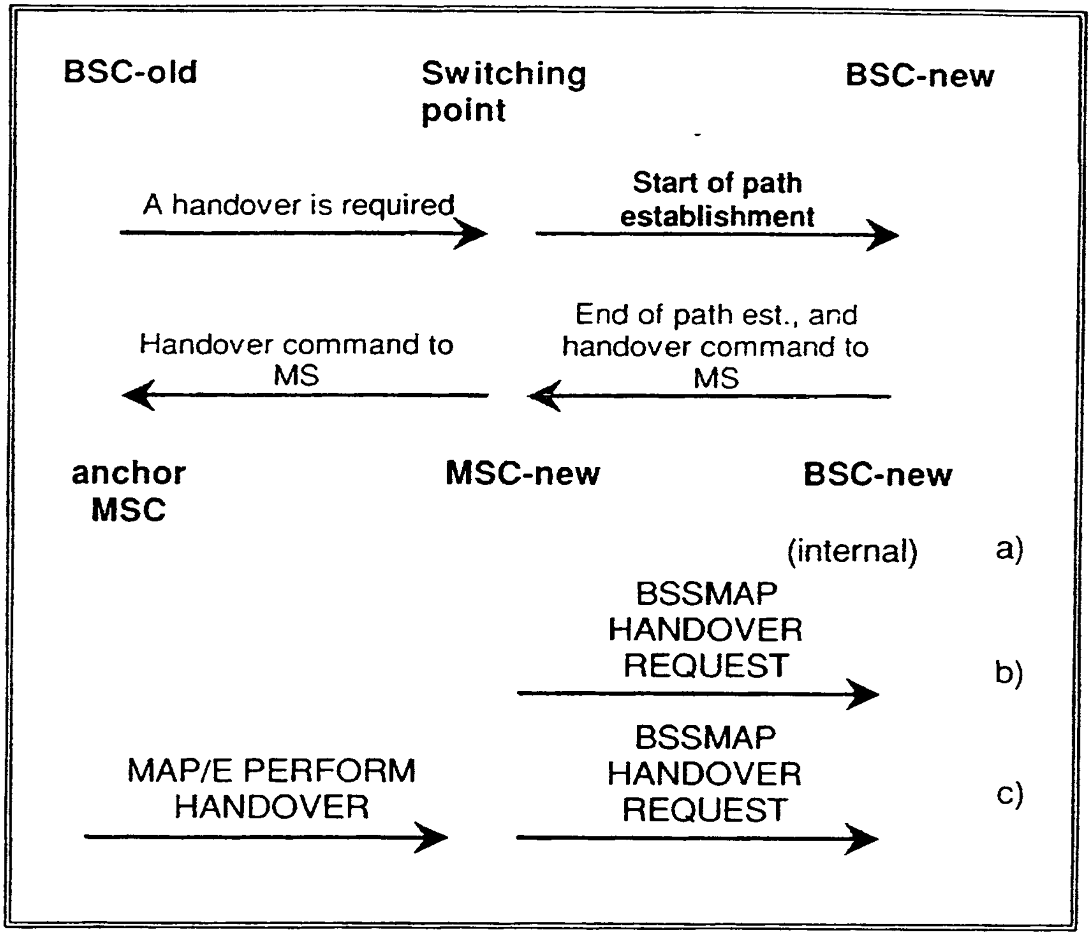


Figure 6.32 – Start of path establishment at handover

The new path is established starting if applicable with the channel on the new A interface. It is triggered by a message coming from the anchor MSC. Only in the next step will the actual circuit between anchor and relay MSC (MSC-new) be established if applicable.

The anchor MSC provides the required information to MSC-new through a MAP/E PERFORM HANDOVER message; when receiving it, MSC-new establishes an SCCP connection with BSC-new, allocates if need be a circuit on the A interface and transmits a BSSMAP HANDOVER REQUEST message to BSC-new, containing the information received in the MAP/E PERFORM HANDOVER message, as explained in case b.

From BSC-new Back to the Switching Point

At this point BSC-new must try to allocate the radio channel. This procedure results in either a positive or a negative answer. Unless

otherwise specified, queuing should not be applied, because other machines are waiting for the answer, and timers are running. Negative cases will be dealt with further on. As for the "happy-ending" case, when a channel can be activated and the corresponding device in BTS-new is ready for mobile station access (an exchange of RSM CHANNEL ACTIVATION and RSM CHANNEL ACTIVATION ACKNOWLEDGE takes place on the (new) Abis interface), BSC-new builds up the RIL3-RR HANDOVER COMMAND message and transmits it to the mobile station, via the switching point and the old resources. It should be stressed that it is BSC-new which builds this message (which will be eventually sent by BSC-old), and thus decides for instance whether the handover will be synchronous or asynchronous, chooses the handover reference, and the initial MS transmission power. In fact, one may consider that BSC-new is in charge of the mobile station from this very moment.

There again, different cases arise depending on the respective positions of BSC-new and the switching point, as shown in figure 6.33.

a. BSC-new is the switching point:

at this point in time, both terrestrial paths are set-up, towards the old and the new BTSs.

b. MSC-new is the switching point:

BSC-new encapsulates the ril3-rr handover command message in a bssmap handover request acknowledge message. Nothing else needs to be done at this stage, since the terrestrial path is already completely established.

c. The anchor MSC is the switching point, and differs from MSC-new:

BSC-new acts as in case *b* above. When receiving the BSSMAP HANDOVER REQUEST ACKNOWLEDGE message, MSC-new inserts the included RIL3-RR HANDOVER COMMAND message in a new envelope, the MAP/E PERFORM HANDOVER RESULT message. This message contains a telephony-like number (provided by MSC-new) to allow the anchor MSC to set up a circuit through normal ISUP or TUP means. This **handover number** is allocated solely for the anchor MSC to establish the circuit with MSC-new, and serves as a reference for MSC-new to link the context with the incoming circuit. The same MAP/E exchange (PERFORM HANDOVER and its RESULT) serves both purposes of providing the information needed for circuit establishment and carrying the RIL3-RR HANDOVER COMMAND message back, ready to be sent to the mobile station along the old path. Upon receipt of the MAP/E PERFORM HANDOVER RESULT message, the anchor MSC is able to

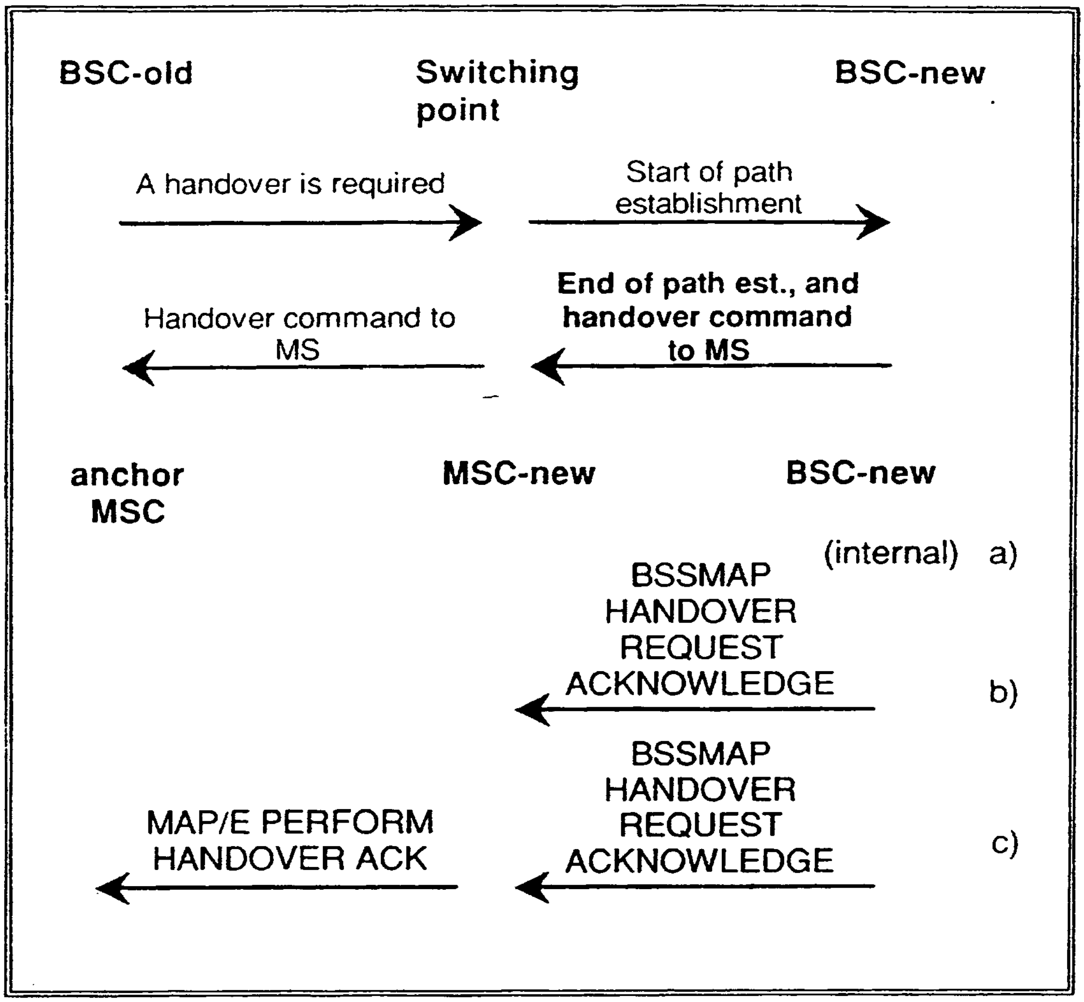


Figure 6.33 – End of path establishment at handover

Once BSC-new has established the new radio channel
(and corresponding resources in the new BTS and on the new Abis interface),
it acknowledges the BSSMAP HANDOVER REQUEST message,
which then triggers in turn and if applicable the set-up of the circuit
between the anchor MSC and the new relay MSC when
the MAP/E PERFORM HANDOVER RESULT message has been received.

set-up the communication with MSC-new, through, e.g., the IAM and ACM messages of ISUP.

Depending upon implementation choices for the switching machine, the paths can be at this stage linked in a two-way conference bridge, a one-way conference bridge or not linked at all. In the first two cases, downlink user data is transmitted to the mobile station via both

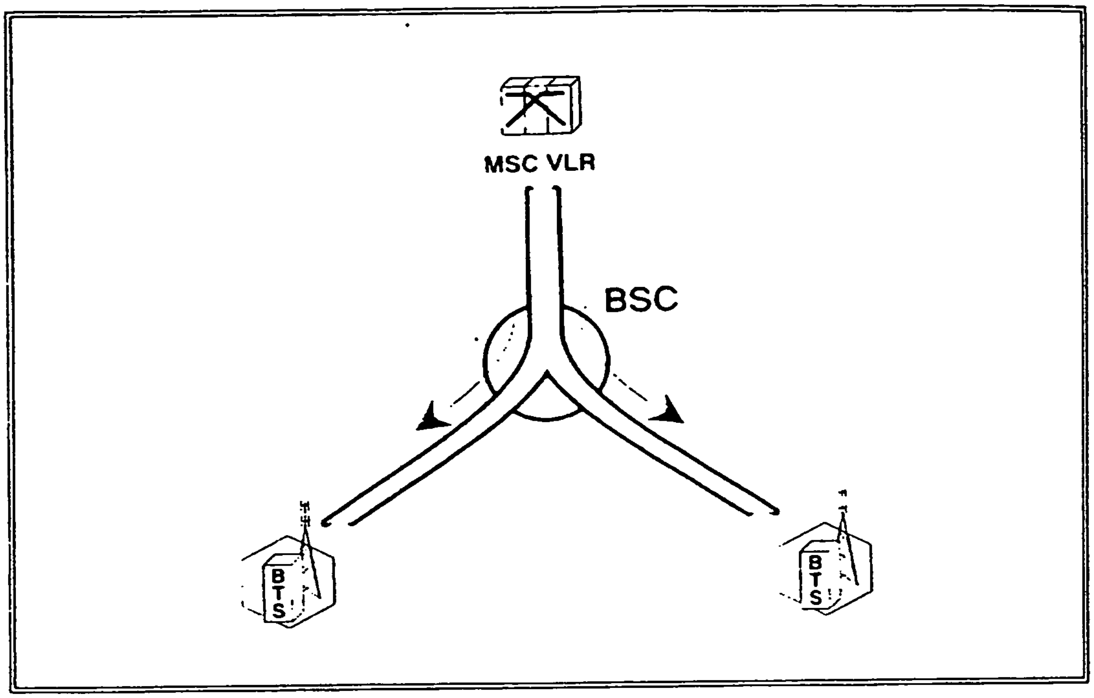


Figure 6.34 - Conference bridge in the BSC for handover

Handover performance can be improved by the insertion at the BSC of a conference bridge, either one-way (downlink) or two-way, so that both paths of an intra-BSC handover may be connected in parallel to the same connection towards the MSC.

BTSs. In the first case only, the two uplink flows are combined into a single one towards the other correspondent (see figure 6.34). It is feasible to insert such a conference bridge only in some cases, for instance when the transmission mode is speech and the transcoders on the old and new paths are different (i.e., speech is carried at 64 kbit/s at the switching point).

From the Switching Point to the Mobile Station

The last step of the first phase of handover execution consists simply in sending the RIL3-RR HANDOVER COMMAND message to the mobile station, as shown in figure 6.35, according to the three following cases:

a. BSC-old is the switching point (BSC-old = BSC-new);
b. MSC-old is the switching point (MSC-old = anchor MSC);
c. the anchor MSC is the switching point, and differs from MSC-old.

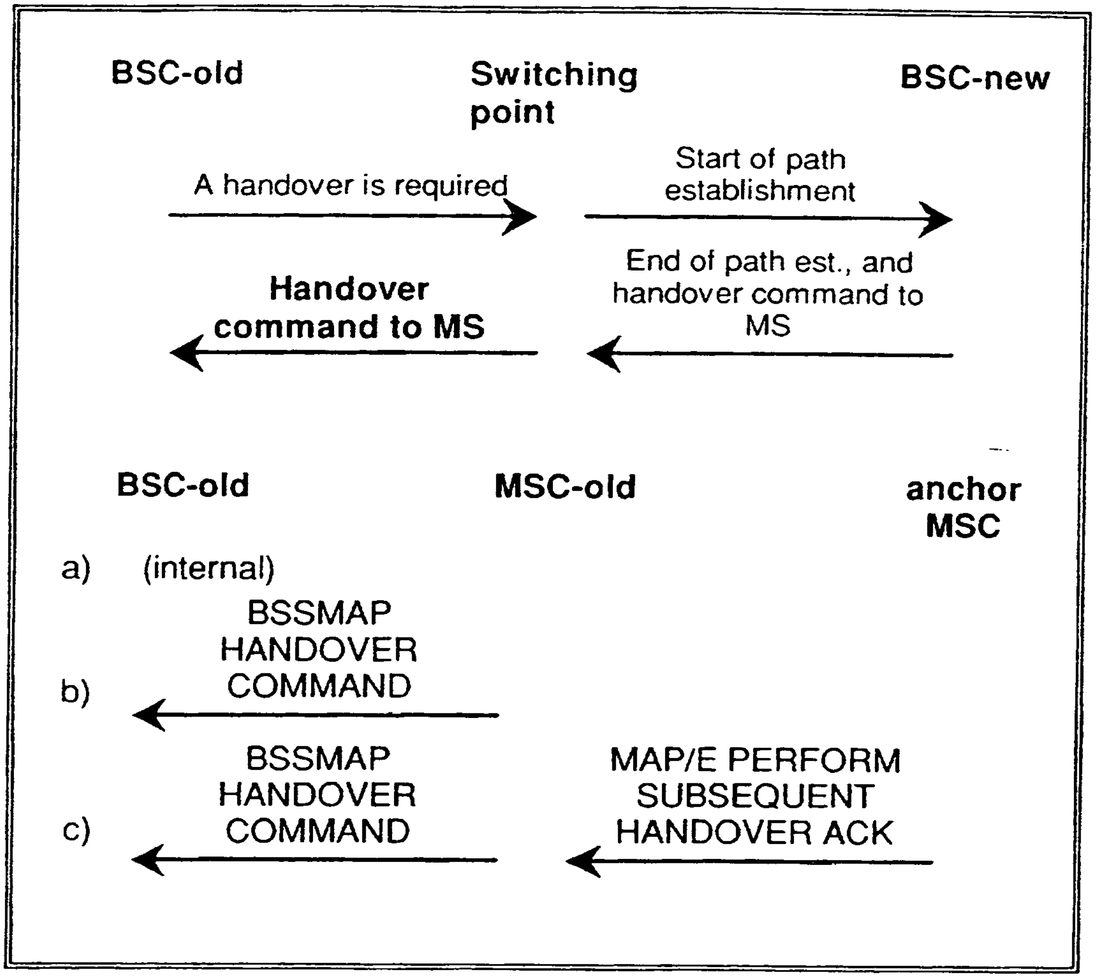


Figure 6.35 – Sending back of the HANDOVER COMMAND

As a last step in the network before mobile station access on the new channel, the RIL3-RR HANDOVER COMMAND message is sent to the mobile station.

The RIL3-RR HANDOVER COMMAND message is carried over the different interfaces in a variety of different envelopes, as shown in table 6.8, which also summarises the transfers already mentioned in the previous paragraphs.

The RIL3-RR HANDOVER COMMAND message identifies the new cell only through its beacon frequency and its BSIC. This is sufficient for the mobile station, and the full cell identity will be read later by the mobile station on the SACCH sent by BTS-new.

Interface	Encapsulating message
between BSC-new and MSC-new	BSSMAP HANDOVER REQUEST ACKNOWLEDGE
between MSC-new and anchor MSC	MAP/E PERFORM HANDOVER RESULT
between anchor MSC and MSC-old	MAP/E PERFORM SUBSEQUENT HANDOVER RESULT
between MSC-old and BSC-old	BSSMAP HANDOVER COMMAND

Table 6.8 – The transfer of the RIL3-RR HANDOVER COMMAND message

The message aimed at the mobile station contains everything the mobile station may need to access the new channel, and is carried unaltered in a variety of encapsulating messages over the terrestrial interfaces.

Not so Successful Alternatives

A number of obstacles may block the smooth succession of events as described above. The main obstacle is the non-availability of radio or terrestrial resources. In this case, an unsuccessful indication is carried back from BSC-new. Considering the longest path, a BSSMAP HANDOVER FAILURE message is transmitted from BSC-new to MSC-new, which in turns triggers backward messages as shown in figure 6.36. Alternatively, a watchdog timer may expire in BSC-old, resulting in a similar situation.

Two possibilities can be envisaged. Either a new handover attempt towards the same cell is performed after some time, or a handover towards another cell is attempted. In the first case, the failure indication goes all the way back to BSC-old, which will re-initiate the handover process when it decides so. All resources which have been allocated along the new path are released. The second case can also be treated this way. However, an alternative exists in the *Specifications*, whereby BSC-old may provide MSC-old with an ordered list of suitable target cells (this applies as already mentioned mainly in the case of rescue handovers, when staying in the old cell is not going to be a good alternative to the first target cell). This list is not conveyed to the anchor MSC if MSC-old differs from it and is the switching point. The failure indication, when reaching MSC-old, will therefore trigger a handover attempt towards the next cell in the list. Only when all cells have been tried in vain will BSC-old be given the failure indication (and the full control of the connection back). This possibility of multiple cell choice in the BSSMAP HANDOVER REQUIRED message is an option for the BSC. Speed requirements tend to favour the multi-cell approach, whereas the optimisation of cell allocation favours the single cell approach, since only BSC-old is in a position to change the cell list according to up-to-date measurements.

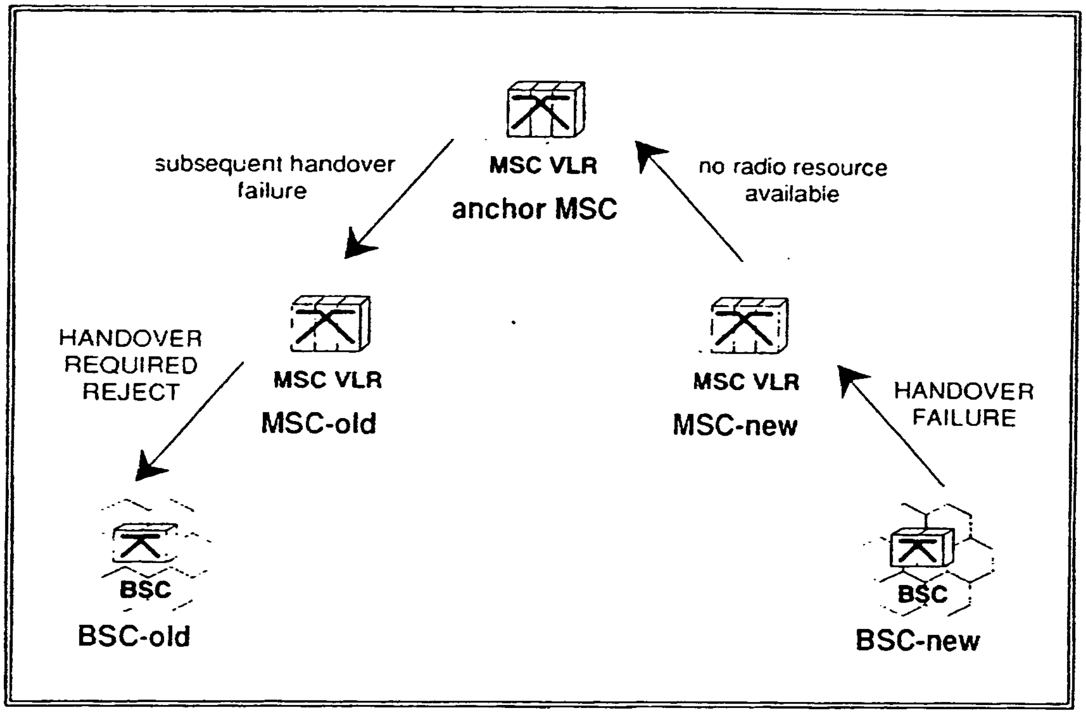


Figure 6.36 – Handover failure at the new BSC

The failure indication goes all the way back to BSC-old as shown, so that the decision to retry or perform another action can be made. Alternatively, MSC-old may attempt handovers toward other cells if it is in possession of a list of candidate target cells.

6.3.4.2. Mobile Station Access and the Conclusion of the Procedure

The mobile station is completely unaware of the infrastructure processes and decisions until it receives the RIL3-RR HANDOVER COMMAND message. As already mentioned, this message contains all the information characterising transmission on the new channel (except for the cipher mode which is assumed to remain the same as on the old channel), and the data needed for the access procedure. In particular, it indicates to the mobile station whether the synchronous or asynchronous handover procedure should be followed. In both cases, thanks to pre-synchronisation, the mobile station is able to synchronise itself quickly on the new channel and starts reception immediately. It will actually receive speech or data from this point, if applicable and if the switching point uses a conference bridge.

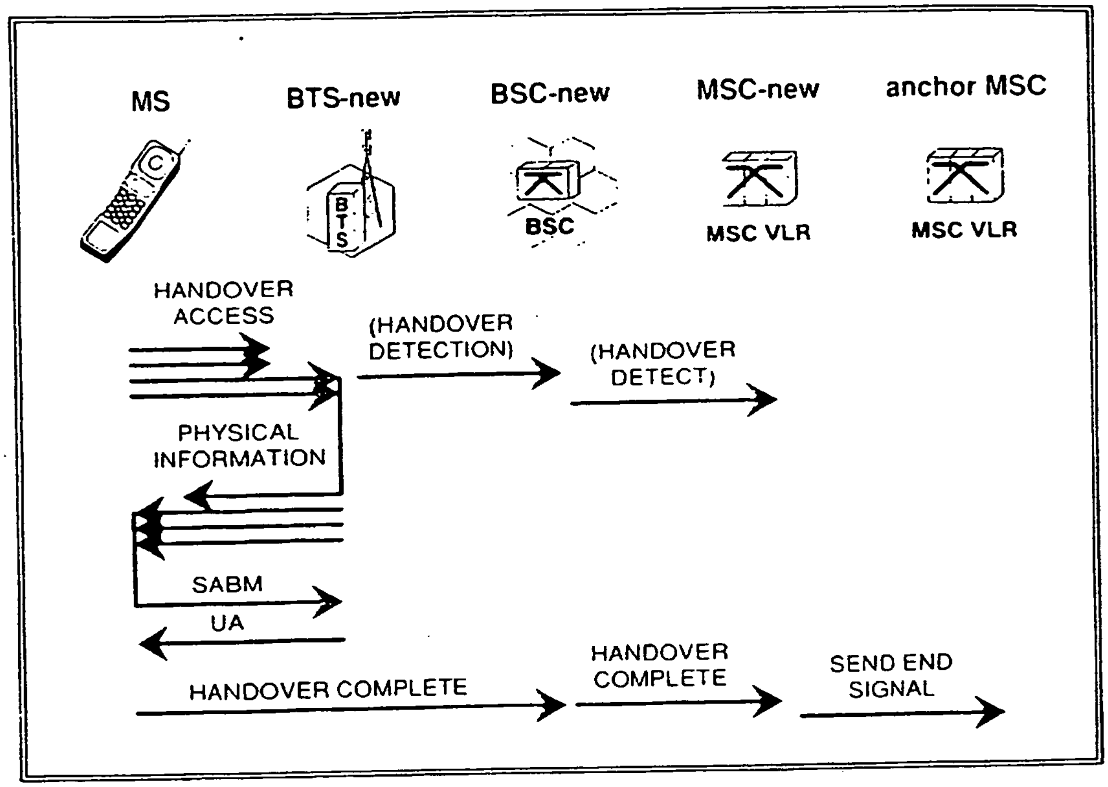


Figure 6.37 – Access in the case of an asynchronous handover

Only following reception of RIL3-RR PHYSICAL INFORMATION messages
does the mobile station switch to normal transmission mode
with the timing advance as indicated.
and sends an RIL3-RR HANDOVER COMPLETE message after having
established the SAPI 0 signalling link on the new dedicated channel.

As far as transmission from the mobile station is concerned, the type of handover intervenes. In the case of a synchronous handover, the mobile station first sends a few access bursts (the RIL3-RR HANDOVER ACCESS message), then starts normal transmission by applying the computed timing advance. If the handover is an asynchronous one (see figure 6.37), the mobile station continues to send access bursts until it has received an RIL3-RR PHYSICAL INFORMATION message from BTS-new, conveying the actual timing advance to apply. Only then does it start normal transmission. In both cases, the RIL3-RR HANDOVER ACCESS message only contains an 8-bit handover reference. This message is the only case where short access bursts are used on a dedicated channel. The handover reference (not to be confused with the handover number) is part of the data transmitted to the mobile station in the RIL3-RR HANDOVER COMMAND message and can be used by BTS-new as an additional check that the accessing mobile station is indeed the expected one.

The RIL3-RR PHYSICAL INFORMATION message is the only case in the *Specifications* where a message above the link layer is sent as an autonomous decision by the BTS. This departure from the general rule is justified by performance requirements. For efficiency reasons, the RIL3-RR PHYSICAL INFORMATION message may be sent several times in a row, until the reception of normal bursts from the mobile station makes it clear to BTS-new that it has received the message. This would not have been so easy if the RIL3-RR PHYSICAL INFORMATION message had been sent by the BSC.

The BTS may as an option indicate to the BSC that it has received adequate RIL3-RR HANDOVER ACCESS bursts on the allocated channel, through an RSM HANDOVER DETECTION message; BSC-new may in turn pass the indication on to MSC-new through a BSSMAP HANDOVER DETECT message. This mechanism allows the switching point (except in the case when it is the anchor MSC, because the information is not carried by the MAP/E protocol) to switch the communication path at this moment without waiting for the full completion of the procedure.

When it is in normal transmission mode, the mobile station sets the link layer to acknowledged mode for signalling messages by sending an SABM frame answered by a UA frame. The mobile station then sends an RIL3-RR HANDOVER COMPLETE message, which will be carried by the infrastructure up to the switching point, when applicable through a BSSMAP HANDOVER COMPLETE message from BSC-new to MSC-new and through a MAP/E SEND END SIGNAL from MSC-new to anchor MSC. The switching point will release the previous path by sending appropriate messages (MAP/E SEND END SIGNAL RESULT from anchor MSC to MSC-old, and BSSMAP CLEAR COMMAND from MSC-old to BSC-old), relayed up to BSC-old which releases the previous radio channel held up until this point. The release of resources which then takes place on the A interface, the Abis interface and at the BTS have no specific differences compared to an RR-session termination.

The sending of the handover complete indication triggers the switching of paths between the old and the new one, if this has not already been done (e.g., upon access detection on the new channel). The question of the necessity of the handover complete indication can be raised at this point: why is there a need of a two-stage mechanism? The difference between the access and the completion is that only the latter triggers the release of the previous channel. Only when it sends the RIL3-RR HANDOVER COMPLETE message does the mobile station abandon all possibility of returning to the old channel. The first stage was added to shorten the interruption time.

The return on the old channel in case of problems is similar to the subsequent assignment case; only the name of the message changes:

RIL3-RR ASSIGNMENT FAILURE in one case, RIL3-RR HANDOVER FAILURE in the other. When this unsuccessful outcome arrives, BSC-old is advised and transmits the information up to the MSC-old if applicable, through a BSSMAP HANDOVER FAILURE message. The MAP/E protocol introduces some limitations. When the anchor MSC is different from MSC-old, there is no means of passing on this information between MSC-old and the anchor MSC, since no message exists on MAP/E for this purpose. The only way for the anchor MSC to react in this case is through timer expiry upon non-reception of a message from BSC-new indicating the completion of the handover. Whatever the means by which it recognised a failure condition, the switching point releases the new path, using normal release procedures, and it is up to BSC-old to decide what action to perform, e.g., make another handover attempt.

An intra-BSC handover is usually performed autonomously by the BSC. As explained in the handover preparation section, it is a BSS implementation option not to involve the MSC (the one in charge of the BSC) at all in the decision when the best cell, as seen from the BSC, is also under the control of the same BSC. In this case, the whole handover will unfold without any knowledge of the MSC. In order to advise it that a handover has occurred successfully, a BSSMAP HANDOVER PERFORMED message is sent from the BSC to the relay MSC, possibly relayed by a MAP/E NOTE INTERNAL HANDOVER message from the relay MSC to the anchor MSC when they differ. This message may also be sent in the case of a handover internal to the relay MSC (case when MSC-new = MSC-old ≠ anchor MSC). The sending or not of this MAP/E message depends on Operation and Maintenance requirements.

6.3.5. Call Re-Establishment

In a radio mobile environment, there is always some risk that a connection will be suddenly cut. This may happen because of a brutal propagation loss, due to obstacles such as bridges, tunnels, or simply buildings in the case of handhelds. But another cell could often be used to continue the communication either immediately or after a very short time (think about a short tunnel through a hill).

The handover preparation and execution are a means of limiting the occurrences of call loss, but they cannot suppress them totally. In the future, when cells become smaller and smaller, the risks will increase. The performance achieved by handover algorithms running on the network side will then lessen, in prediction accuracy as well as in reaction time, and connection loss probability will increase. The mobile station has in fact some ways to determine that a handover is needed, and may be

more efficient in these cases. In some systems crafted for a microcell or pico-cell environment, all handovers are triggered by the mobile station alone. However, network handover control has many advantages when it can be applied: this stems from the obvious fact that the network has a much better understanding of the general situation than any single mobile station. Call re-establishment may then be considered as a kind of mobile station triggered handover, but limited to the extreme case of rescue handover when communication with the current cell is effectively lost. One could foresee that the importance of this feature will grow in the future. For instance, it can be imagined that in some environments the procedure will be triggered sooner, so as to improve the system performances where network triggered handover will have shown its limits.

Despite these considerations, the call re-establishment procedure is a poor relative in the GSM procedure tribe, and has serious limitations in phase 1. Let us see what there is of it.

As a general point, it should be noted that the call re-establishment procedure is not a full-blown RR procedure. We will see in this section messages from the RIL3-MM protocol. Only because of its kinship with handover do we present the re-establishment procedure in this chapter and not in Chapter 7.

Call re-establishment has two parts. The mobile station has the leading role in the first one, which is very close to an access procedure. The second part is the network's, and consists in the recovering of upper layer contexts.

The closeness of call re-establishment with initial access is quite normal, because the mobile station has to start from scratch. The differences are important, and come mainly from the requirement for speed: when a connection is lost, a timer starts ticking in the anchor MSC, and at its expiry everything related to the moribund connection is erased. As a consequence, any fraction of second lost in the call re-establishment procedure increases the risk of total loss.

The first issue is to determine the new cell. The speed requirement limits the choice to the neighbouring cells already known and with which the mobile station is pre-synchronised, because finding new cells and getting synchronisation may take seconds. The selection rule is to just select the one with the highest signal strength. The parameters to compute the different radio criteria (the $C1$'s of the possible cells, which will be described in Chapter 7) as for idle-mode cell selection are not known, and receiving them takes time. The *Specifications* nevertheless require the mobile station to check the radio criterion for the chosen cell. This constraint forces the mobile station to wait for the reception of a BCCH

message containing the required parameters, and this can take up to three quarters of a second (if the first decoding attempt succeeds!). Another formality has to be checked: the chosen cell must not be barred, and call re-establishment must be allowed in it. The corresponding indications are part of all BCCH messages, so that this checking does not add any more delay.

After having received the required BCCH message, and checked the radio criteria and authorisations, the mobile station sends an access request on the RACH. Though not specified in the *Specifications* (as many other small details of the procedure), it seems that it is allowed to use the RACH on TN 0, even if there are others. The access request indicates the reason for access (i.e., call re-establishment), so that the network is aware of the critical nature of the request. The required channel is not indicated, but the network can easily play safe and allocate a TACH/F.

The initial message is an RIL3-MM CM RE-ESTABLISHMENT REQUEST. Its information contents are minimal: the subscriber identity and the classmark. The mobile station does not volunteer anything else, and the network has to use this to find out everything about the lost connection! Among the conspicuous missing data known directly or indirectly by the mobile station are the cell with which the connection was lost, the identity of the anchor MSC and the required type and mode of the channel.

In any case, the only case catered for by phase 1 protocols is when the new cell and the previous ones are managed by the same MSC. The anchor MSC is then implicitly determined. Moreover, even if the previous cell was known, it would be to no avail since no mechanisms have been included to recover the RR-session except when the new MSC is the anchor MSC. An inference is that call re-establishment is impossible when there is a relay MSC.

The fact that the required type of channel is not given by the mobile station has no clear explanation. This is a source of delay, because the BSC has to wait for the indication from the MSC to allocate the right type of channel. Unless, as already mentioned, if the BSC gambles and initially allocates a TACH/F.

The recovery of the contexts is then entirely an MSC issue, and must be done with only the subscriber identity to start with. From this the MSC must find the old context (if it still exists—it could have been erased after a timer expiry, or simply because the correspondent was not patient enough). Then the MSC performs an assignment procedure and possibly a ciphering start procedure, telling the BSC the type of channel required, the mode, and so on, and allocating the BSC-MSC terrestrial

route when need be. The MSC may even choose to perform an authentication despite the incurred delay. The BSC then performs the needed procedures within the BTS and with the mobile station (subsequent assignment, ciphering start, mode modification, ...). Only then can an RIL3-MM CM SERVICE ACCEPT message be sent to the mobile station, and the end to end communication resumed. Let us note that means for the network to reject the request have been foreseen. The RIL3-MM CM SERVICE REJECT serves this purpose (in particular with the cause "call cannot be identified", which is likely to be used a lot).

A number of other fuzzy points exist in the *Specifications*. They include in particular, the recovery of several CM-transactions and the corresponding transaction identifiers (see Chapters 5 and 8), and, more important, all the cases of collision, when the connection loss happened during an ongoing procedure in any layer. Another issue is the release of the old path. As far as can be analysed, the anchor MSC can and should release it once aware of the re-establishment attempt, even if in the same BSC (or the same cell).

One wonders if call re-establishment will really be used in phase 1. This is hopefully an area in which the future phases of the *Specifications* will bring improvements.

6.3.6. RR-SESSION RELEASE

When all needs for an RR-session have disappeared, for instance because a location updating procedure has ended, because a call is terminated, or because of a failure, the mobile station must go back to idle mode and the resources must be released, in order to be free for allocation for other needs. This release mechanism is done through a so-called "normal release" procedure, which is always triggered by the anchor MSC.

If distinct from the relay MSC, the anchor MSC releases the RR-session by sending a MAP/E SEND END SIGNAL RESULT to the relay MSC on one hand, and by releasing the circuit if present, through ISUP release procedures.

The next step is the BSSMAP CLEAR COMMAND message sent from relay MSC to BSC. This message may be piggybacked on an SCCP RELEASE message releasing the SCCP connection. In this case, the BSC acknowledgement (BSSMAP CLEAR COMPLETE message) must be piggybacked on the SCCP RELEASE COMPLETE message. The clearing actions of the BSC can take place in parallel with the sending of the BSSMAP CLEAR COMPLETE message, since the *Specifications* do not

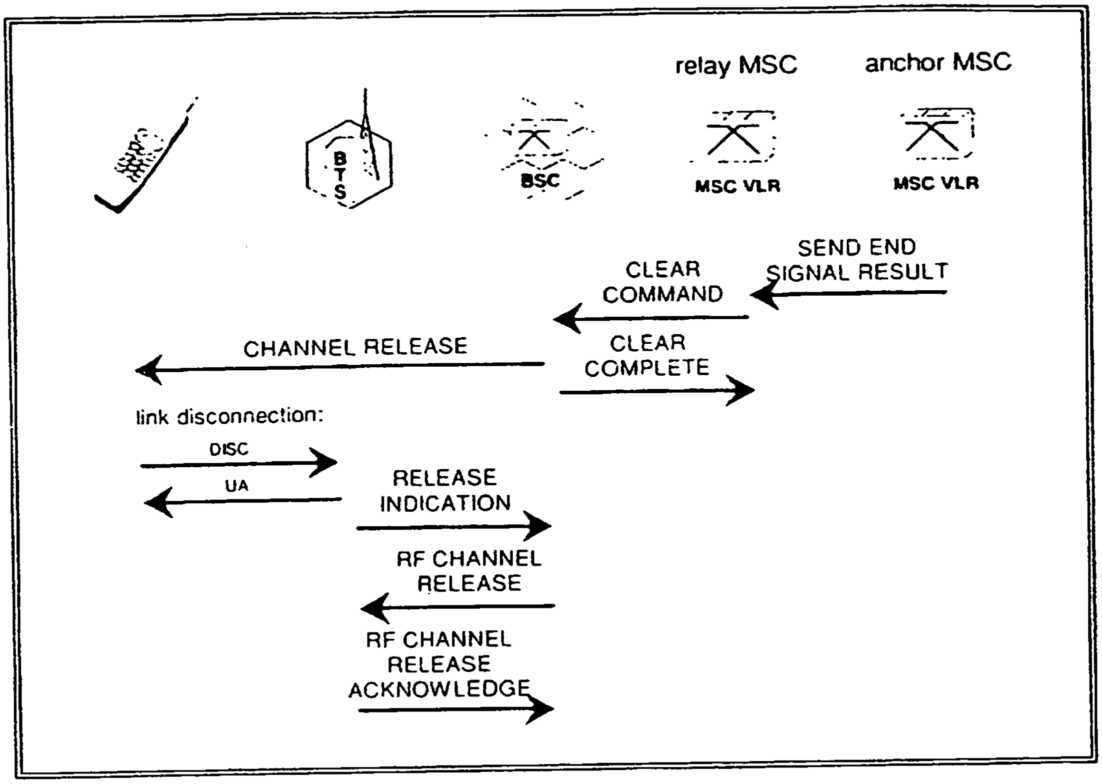


Figure 6.38 – The normal release procedure of an RR-session

Normal release is always triggered by the anchor MSC, but the BSC manages the return of the mobile station to idle mode before releasing the BSS resources.

impose any specific order between these two actions. Once the BSC has ordered the mobile station to go back to idle mode through an RIL3-RR CHANNEL RELEASE message, the mobile station disconnects the signalling link, and this event is reported by the BTS to the BSC through the RSM RELEASE INDICATION message. The *Specifications* include a number of timers and repetitions in order to ensure that whatever frame losses may occur during this period the mobile station eventually goes back to idle mode and stops using the channels. This is of prime importance to avoid allocating a channel to a mobile station when another mobile station may still transmit on this same channel. Only when the BSC is sure that the mobile station has left will it de-activate the BTS device, through the RSM RF CHANNEL RELEASE / RSM RF CHANNEL RELEASE ACKNOWLEDGE exchange. The corresponding radio channel is then considered as part of the pool of free channels by the BSC. The whole procedure is illustrated in figure 6.38.

An RR-session may also be released in other conditions, for instance when the infrastructure has lost actual contact with the mobile station. Such a situation may arise when propagation conditions are bad,

or when the interference level is too high. One role of handover is to cope with such cases, but this is not always possible, e.g., when the user has gone out of coverage (in an underground car park for instance), or switched off his mobile station in the middle of an RR-session. Such cases must be detected, in order for the infrastructure to free the corresponding resources.

The mechanism specified in the *Specifications* for this purpose consists in having both the mobile station and the BTS monitor the message loss rate on the SACCH. Let us recall that messages are sent regularly (about twice per second) in both directions of the SACCH, throughout the life of a connection. A powerful error detection mechanism has been included in signalling messages, enabling the receiver to estimate message loss. This estimation is done through a counter, incremented in case of a correctly received block and decremented in the other case (see figure 6.39). When the counter reaches a minimum threshold, the link is considered as broken. On the mobile station side, this event leads to a return to idle mode (the mobile station may subsequently attempt a re-establishment, see page 412). The infrastructure is able to adjust the mobile station behaviour, in order to allow fine tuning on a cell basis, although it seems unlikely that it will be necessary to manage differences between mobile stations. The relevant parameters are sent regularly on the SACCH as well as on the BCCH. Setting them to satisfying values must be done through field experiments in operational networks. On the infrastructure side, the counter is in the BTS and the failure indication is given to the BSC in an RSM CONNECTION FAILURE INDICATION message.

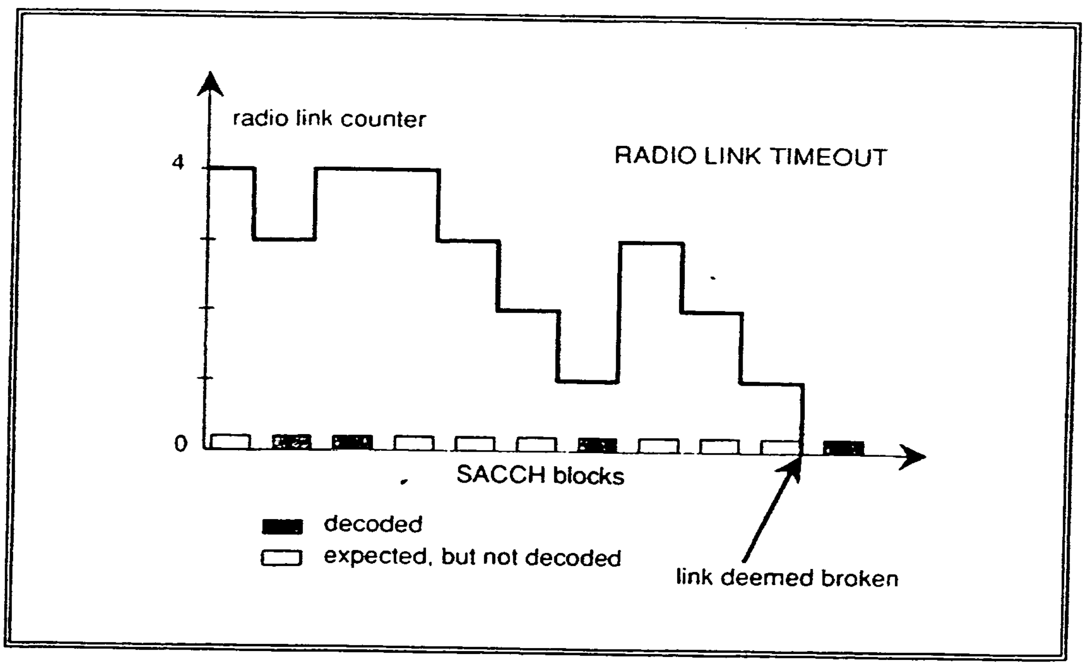


Figure 6.39 – SACCH counter for link management

A counter enables each receiver to estimate the frame loss rate on the SACCH (downlink for the mobile station, uplink for the infrastructure).
When this counter reaches 0,
the link is deemed broken and actions are taken to release the resources.

While the detection of transmission loss simply triggers the mobile station to abandon the connection and return to idle mode, things are a little bit more complex on the infrastructure side. Since both directions of transmission may experience different qualities, reaching the threshold in the uplink direction does not necessarily imply that the mobile station also experiences a break of the link. The infrastructure must nevertheless make sure that the mobile station leaves the channel before deeming the channel to be free. The BSC therefore commands the BTS to stop transmission of downlink SACCH frames (by an RSM DEACTIVATE SACCH message), so that the mobile station counter will inexorably reach the minimum threshold after some given time. The monitoring of the uplink channel may take place either in the BTS or in the BSC. In the first case, the BTS may report the loss of the link through an RSM CONNECTION FAILURE INDICATION message to the BSC.

Once the link failure is detected and notified to the relay MSC through a BSSMAP CLEAR REQUEST message, the same A interface exchange takes place as in the case of normal connection release, i.e., the MSC sends a BSSMAP CLEAR COMMAND message, answered as expected by a BSSMAP CLEAR COMPLETE message from the BSC.

6.3.7. LOAD MANAGEMENT PROCEDURES

A few procedures in the RR plane allow the MSC and the BSC to deal with overload situations. They include means to exchange information between machines so that each one gets the information it needs about the current load situation; and means to act so as to limit the effect of the overload. Procedures dealing with load management appear in two main areas: RACH and PAGCH load, and TCH load.

6.3.7.1. Load on Common Channels

Some information concerning the load on the RACH and on the PAGCH could be inferred by the BSC from the requests it sends to, or receives from, the BTS. Still, the BTS is in a better position to assess the exact load on these channels.

A message, RSM CCCH LOAD INDICATION, has thus been introduced in the RSM protocol, to enable the BTS to send some information about the RACH and PAGCH loads to the BSC. The conditions for sending this message are set through the Operation Sub-System; it can be regularly sent, or only when the load on one of the channels is above some threshold. The message pertains to a single pair RACH/PAGCH.

The BSC may use this information to change the RACH load control parameters included in the corresponding BCCH, and to modify its assignment priority rules.

6.3.7.2. Load on Traffic Channels

The number of dedicated channels of each sort currently allocated in a cell is known by the BSC. This information may also be useful to the MSC, for instance to balance the traffic between cells. The number of TACH/F (and in phase 2 of TACH/H) currently allocated can be indicated to the MSC with a BSSMAP RESOURCE INDICATION message. The message can be sent in various cases, which are controlled by the MSC and indicated with a BSSMAP RESOURCE REQUEST message. The MSC can ask for one single immediate message; for regular sending; or for spontaneous sending when some conditions are met (these conditions are set through the OSS).

The MSC can do many things to limit useless signalling towards the BSC. However, the resource indication function was introduced mainly to support traffic handovers under the control of the MSC. In case of an overload condition local to a cell, handovers can be used to balance the traffic between cells. This is the purpose of the "handover candidate enquiry" procedure, which is used when the overloaded cell has neighbours under another BSC. Only the (relay) MSC is in a position to know which of the neighbour cells can take over a part of the load. By sending a BSSMAP HANDOVER CANDIDATE ENQUIRY message, the MSC indicates to the BSC that it would be better for traffic balance to hand over a given number of connections from one cell to other cells in a given list. When the BSC is done with the consequences of this message (precisely after having sent one BSSMAP HANDOVER REQUIRED message, with the appropriate cause, for each of a number of connections), it sends a BSSMAP CANDIDATE RESPONSE message. What the BSC must do is not specified in detail. As already mentioned, traffic balancing is contradictory with the choice of the cell for spectral efficiency reasons, and the action of the BSC must be designed with care, to avoid having the handed-over connections handed back a few seconds later because of radio criteria. It may be noted that an alternative strategy to solve traffic imbalance is to use the OSS to control handovers directly via the BSC. This uses the parameter modification procedures discussed in Chapter 9.

6.3.7.3. General Overload

In addition to the specialised procedures described above, the BSSMAP and the RSM contain a number of procedures to cope with overload in general, whether of transmission resources or computation capacity. The *Specifications* describe the messages conveying the information from MSC to BSC, from BSC to MSC (BSSMAP OVERLOAD and from BTS to BSC (RSM OVERLOAD) that the recipient must "reduce the traffic".

The purpose of overload control is to reduce the traffic as close to its source as possible. Let us examine the different cases:

- BTS to BSC: It is not clear how a BTS can be overloaded otherwise than on the common channels (and this is dealt with by specific procedures), since a BTS is normally designed to cope with the simultaneous usage of all its radio channels. The only thing it seems the BSC could do in answer to an RSM OVERLOAD message is to consider that only a portion of the channels can be used;
- BSC to MSC: The only traffic on the A interface established under the control of the MSC concerns the mobile terminating calls. A possible reaction of the MSC to the BSSMAP OVERLOAD message is then to reject a portion of the mobile terminating calls rather than to send the paging messages. On the other hand, this can be done by the BSC, and with a better correlation with the congested cells;
- MSC to BSC: An MSC overload indication can be used by the BSC to reduce the number of mobile station accesses it accepts, by using one of the various means described on page 370.

6.3.8. SACCH PROCEDURES

When the mobile station is in dedicated mode, it is always allocated a bi-directional channel of limited capacity (the SACCH), in addition to the main channel conveying the information for which the connection exists. The SACCH is used for a variety of functions. The main item, which justifies the existence of the SACCH, concerns the continuous monitoring of the connection in a mobile environment: transmission power control, timing advance control and measurement reporting. The second role of the SACCH is to convey general information to the mobile station.

6.3.8.1. Radio Transmission Control

In the downlink direction, the SACCH carries the commands related to power control and timing advance. These commands from the

network are carried in the so-called *L1-header*, or layer 1 header, meaning that this information pertains to the physical layer in the radio path protocol architecture, and is therefore formatted outside the scope of the link layer, independently from the messages carried inside link layer frames. The required power control level and timing advance are then sent once per SACCH message, i.e., about twice per second. In the uplink direction, the same header exists, and contains the corresponding "acknowledgement" by the mobile station. It is coded in a similar way, and includes the values of the two parameters in use at the end of the preceding measurement period. The actual timing advance used should be equal to the one ordered, whereas the actual power level may differ from the one ordered because of the maximum variation speed.

The timing advance is managed autonomously by the BTS. On the other hand, the transmission power is basically controlled by the BSC. The BSC analyses the measurements, and uses the RSM MS POWER CONTROL and RSM BS POWER CONTROL messages to convey the requirements to the BTS.

Measurement reports are sent by the mobile station on the uplink SACCH at every possible opportunity, and at least once per second. The exact specification requires that, among any two successive messages on the uplink SACCH, at least one be an RIL3-RR MEASUREMENT REPORT message. In a basic scheme, the BTS generates a message toward the BSC (RSM MEASUREMENT RESULT) at every measurement period (about twice per second): this message indicates whether or not a measurement message was received from the mobile station, conveys its contents in the first case, and includes the result of the measurements performed by the BTS. All these data are then processed by the BSC for transmission power control and handover preparation.

The schemes described above are completely defined in the *Specifications*. However, they present a severe defect: they result in a very substantial signalling load on the Abis interface. Hooks have been introduced in the protocol for a scheme where more processing is performed in the BTS (and less in the BSC), so as to reduce the information flow in-between. This alternative scheme is only outlined in the *Specifications*, inasmuch as the messages or message elements that can be used are not completely specified (RSM PRE-PROCESS CONFIGURE, RSM PRE-PROCESSED MEASUREMENT RESULT, RSM PHYSICAL CONTEXT REQUEST and RSM PHYSICAL CONTEXT CONFIRM messages, and *MS POWER CONTROL PARAMETERS* and *BS POWER CONTROL PARAMETERS* information elements). In order for these messages to be used, complementary specifications are needed from the operator or the manufacturer. Within these constraints, transmission power control can be entirely taken in charge by the BTS, possibly with some parameters

(e.g., maximum power) set up and modified by the BSC. Handover preparation is a more tricky business. The pre-processing of the measurements in the BTS can be anything from none (the basic scheme) up to include the decision that to trigger a handover is necessary or useful, barring load considerations. In the latter case, the BSC would intervene only when necessary, to check the load and channel availability aspects and to trigger effectively the handover.

6.3.8.2. General Information

A second use of the SACCH is the transmission of general information from the network to the mobile station. This information includes parameters specific to the radio connection, but not to upper layers, and which are not so important as to require the stealing of user information. Basically, the transmission of these parameters is useful only at the beginning of a channel connection (in particular after a handover) and when they change, which is very seldom. However, because there is no high load on the downlink SACCH and because these messages are not acknowledged, they are repeated continuously as a background task. They bear some relationship with the messages broadcast to all mobile stations in a cell on the BCCH (RIL3-RR SYSTEM INFORMATION TYPE 1 to 4 messages), which will be described in a later section, and their names (RIL3-RR SYSTEM INFORMATION TYPE 5 and 6) reflects this similarity. However, their information contents are different; only a part of the information broadcast on the BCCH needs to be sent to a mobile station in dedicated mode. This information includes:

- parameters for monitoring the measurement process (list of frequencies to monitor, BSIC screening, BCCH frequency indication);
- parameters for controlling the radio link failure detection (see page 417);
- requirements for the application of uplink discontinuous transmission; and
- other information which is of no direct relevance in the present state of the *Specifications* (full cell identity).

The bulk of this information concerns the measurement process. It must be noted that, after a handover, the mobile station must await the reception of this information before starting measurement reporting.

The SACCH is always used in both directions for actual transmission, to allow the other side to do reliable measurements and to detect radio link failures. This is the reason of the continuous sending of

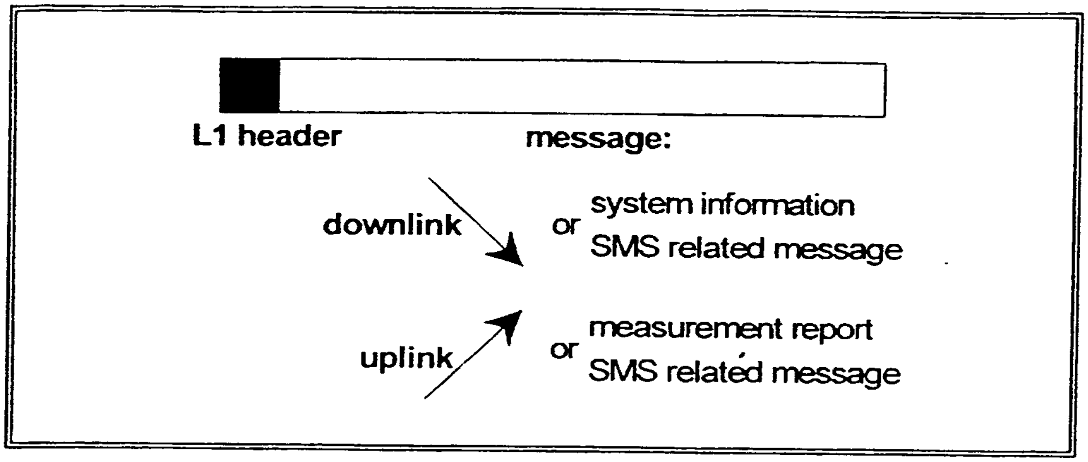


Figure 6.40 – Contents of an SACCH block

Except for messages related to the transfer of short messages, the basic use of the SACCH concerns the reporting of measurements from MS to BTS, and the sending of general information from BTS to MS (the RIL3-RR SYSTEM INFORMATION TYPE 5 or 6 messages).

RIL3-RR SYSTEM INFORMATION messages downlink and RIL3-RR MEASUREMENT REPORT messages uplink, when nothing else needs to be sent.

A last use of the SACCH consists in carrying short messages when the main channel is a TCH/F. This capability allows the mobile station to either send or receive short messages while engaged in a communication on the main channel. This situation represents in fact the only case in phase 1 where two independent telecommunication services can be provided in parallel.

As far as short messages are concerned, the functions of the BTS are limited to the management of the physical layer and of the link layer described in Chapter 5. The actual decision to send a short message is taken by other network entities. For the sending of general information, on the other hand, it would not be efficient to ask a machine other than the BTS to decide on each individual sending of an RIL3-RR SYSTEM INFORMATION message. The BTS is then in charge of the repetition of these messages on the SACCHs. The BSC only provides information regarding these messages when their contents are modified, using for this purpose the RSM SACCH FILLING message. The procedure is such that the information is not provided on a connection basis, but on a TRX basis.

Figure 6.40 summarises the different types of information which transit on the SACCH, both uplink and downlink, and which are scheduled in such a way as to achieve 100% use of the channel.

6.3.9. FREQUENCY REDEFINITION

The frequency redefinition procedure is used to change the frequency properties of the frequency-hopping channel used by a mobile station in synchronism with other mobile stations, in order to cope with a change of the frequencies used in the cell.

It is a very simple procedure, at first sight. It consists solely in the sending of one message to each concerned mobile station, the RIL3-RR FREQUENCY REDEFINITION message, as well as one message to each corresponding BTS device. However, there is no RSM message defined for this purpose, and the management of a frequency redefinition between BSC and BTS must use a procedure from the operation and maintenance BSC-BTS protocol for setting RF parameters in the BTS to command the configuration of the BTS. The RIL3-RR FREQUENCY REDEFINITION message includes both a starting time and the new frequency parameters (the list of frequencies as well as the MAIO and HSN, as explained on page 360), and the starting time can be set in the BTS by the operation and maintenance protocol on the Abis interface (the BTS Management protocol, BTSM). This state of affairs spoils the functional split between the two protocols, as a BTSM message is used for real-time management of connections.

6.3.10. GENERAL INFORMATION BROADCASTING

We have seen in this chapter that some information of a technical nature is needed by the mobile stations in *idle mode*, such as the configuration of the common channels. More reasons for the need of such information will be exposed in the next chapter. This data is regularly broadcast on the BCCH. It covers items of various nature which are cited here and there in this chapter and the next one. This section will act as a summary for the description of this information.

The BCCH is a low capacity channel, able to transmit one 23-octet long message every 0.235 second. It is therefore a scarce resource. The repetition rate for the different information items must be chosen as a trade-off between the use of the BCCH resource and the resulting time for the mobile station to get access to the information. As a consequence, several messages have been defined which contain different contents and which have different periodicity.

The broadcast items will now be listed together with their use and recurring rate: their order of description is based on the mobile station phase in which the information is needed.

6.3.10.1. Cell Selection Information

An important part of the broadcast information is related to the cell selection process. Some information of this kind, is also directed towards mobile stations of neighbouring cells. This information will be addressed in Chapter 7, when the cell selection criteria will be described. It includes the location area identity and various parameters impacting the access choice, including the indication whether the cell is barred for access or not. The corresponding parameters in the *Specifications* are called *LAI*, *CELL SELECTION PARAMETERS* and (for the *CELL_BAR_ACCESS* flag) the *RACH CONTROL PARAMETERS*.

The transmission rate has a direct impact on the time it takes a mobile station to include the cell in its comparison list, from the moment when the BCCH is received correctly by this mobile station. This is not such a stringent constraint, but the information is nevertheless broadcast at a high rate: 2 occurrences out of 4, and even more in the case of the *CELL_BAR_ACCESS* indication which is part of every single BCCH message. It is worth noting that, even when several *time slots* are used for the common control channels, mobile stations currently located in neighbour cells only listen to the BCCH on TN 0. Therefore, the rate of transmission on other time slots need not be so frequent. For sake of simplicity, though, the specifications are the same for other time slots as for the prime BCCH.

Another constraint impacts the information used for cell selection. Mobile stations in neighbouring cells must listen to their paging sub-channel. Since recurrence on the PAGCH is a multiple of 51 × 8 BP, care must be taken that the relative position of a paging sub-channel and of the cell selection information in BCCH messages never results in a systematic masking of the latter. This is achieved through two mechanisms:

- the maximum rate of a paging sub-channel is half the rate of a BCCH. Therefore, a paging sub-channel may at worst mask half of the messages of any BCCH, but never all messages;

- the cell selection information is not transmitted in every other message, but according to the following pattern:
  in, in, out, out, in, in, out, out, ...

These configurations result in the masking, at worst, of every second occurrence of the cell selection information by any paging sub-channel.

6.3.10.2. Information for Idle Mode Functions

Information for idle mode functions is used by mobile stations once they have selected the cell and are staying there for some time in idle mode. They include the configuration of the common channels, the neighbour cells to monitor and the configuration for cell broadcast messages. Once settled in the cell, a mobile station will regularly check the values for changes; this checking is done at a very slow rate and the transmission rate therefore only impacts the initial settling time.

A first need concerns the common channel configuration, i.e., the number of *time slots* used for common channels (this is enough for a given mobile station to know which *time slots* are used and which one to listen to), as well as the parameters enabling the mobile station to calculate where to find its own paging sub-channel. This is found in the *CONTROL CHANNEL DESCRIPTION* parameter.

Once camped on a cell, the mobile station must also know the list of beacon frequencies to monitor. This is found in the *NEIGHBOUR CELLS DESCRIPTION* parameter. This element is also relevant for dedicated mode, but only as a default value during the short time between access and the reception of the first SACCH message containing the information. Because of this possible usage for measurement reporting in dedicated mode, this parameter contains a binary indicator (the *BA_IND* flag) enabling the list of beacon frequencies to be changed while keeping the network aware of which one was used by the mobile station in its reported measurements.

Lastly, mobile stations equipped for receiving broadcast short messages must know whether the cell provides a Cell Broadcast CHannel (CBCH), as well as where to find it when applicable. The *CBCH CHANNEL DESCRIPTION* and, if needed, *CBCH "MOBILE ALLOCATION"* information elements are therefore broadcast for this purpose.

Each of these three types of information is broadcast in every fourth message, i.e., roughly once per second.

6.3.10.3. Information Needed for Access

Sooner or later, the mobile station will want to access the cell, i.e., to obtain a bi-directional dedicated channel for its transmission needs. This may happen at the point of entry into the cell, when location updating is required or for a call re-establishment. Critical time constraints therefore apply to some of this information.

As explained in the section dealing with random access (see page 368), some control means are available for the BSC to limit access attempts, through the mechanism of "access classes". The BCCH information therefore includes the list of access classes allowed for access and the indication whether emergency calls are allowed. Another flag indicates whether call re-establishment is allowed. The BSC also controls the scheduling of access attempts and repetitions, and the corresponding parameters are broadcast. All of these parameters are part of the so-called *RACH CONTROL PARAMETERS*.

The most critical case of access to this information is call re-establishment. Any slowing down of the access procedures in this case increase the probability of losing the call. For this reason, all the access control information is broadcast in every BCCH message, i.e., 4 times per second.

Once having sent its RIL3-RR CHANNEL REQUEST burst on the RACH, the mobile station must decode the corresponding channel assignment message sent back by the network, if any. Since such a message must fit in a single block, its length may be too short in case of frequency hopping to send a complete description of the channel frequency characteristics. Therefore, as explained on page 361, part of the channel description information, applicable to *all* initial assignments, is broadcast regularly (once per second), in the *CELL CHANNEL DESCRIPTION* parameter. In most cases, this information will correspond to the list of all frequencies which might be used for dedicated channels in the cell, although it need only contain the list of all frequencies which might be used for *initial* channel assignment in the cell.

This list is critical information for access, since the other frequency parameters of the allocated channel cannot be understood by the mobile station without it. However, a mobile station could perfectly start the access procedure (i.e., send an RIL3-RR CHANNEL REQUEST message) before decoding the corresponding BCCH block, and therefore wait until it has done so after having received the network initial assignment, before actually accessing the dedicated channel. The *Specifications* are not clear on this point, and it seems a good way to improve the performance of call re-establishment.

6.3.10.4. Information for Mobile Stations in Dedicated Mode

Strangely enough, part of the information broadcast on the BCCH has no application other than after access, i.e., for mobile stations in dedicated mode. This information, given again on the SACCH, includes parameters to control the reporting of measurements, in particular the BSIC screening information included in the *PLMN PERMITTED* parameter, to which we will come back in Chapter 7. It also includes the "power control indicator" which we encountered in the section dealing with measurements, as well as the indication whether mobile stations are obliged, forbidden or permitted to use uplink discontinuous transmission. These three indications are included in a *CELL OPTIONS* parameter.

This information bears no real timing constraint; its absence would indeed little affect system performance. It is sent in every fourth message.

6.3.10.5. Cell Identity

The final piece of information found on the BCCH is the complete *CELL IDENTITY*, which is sent in every fourth message. This element has no direct usage as far as the contents of the *Specifications* can be analysed, although it seems mandatory because of general radio regulations. Beside, it may be useful for network testing purposes.

6.3.10.6. Message Scheduling and Contents

The different items described above are grouped in four different messages for GSM900, bearing the non-informative names of SYSTEM INFORMATION TYPE 1 to 4. An additional RIL3-RR SYSTEM INFORMATION TYPE 2BIS has been defined for DCS1800, in order to cope with the extra length of the list of frequencies coming from the number of available frequencies. All these messages are sent according to an 8 × (51 × 8) BP cycle, which includes 8 message occurrences for a duration of about 2 seconds. The scheduling is performed according to the description of table 6.9. RIL3-RR SYSTEM INFORMATION TYPE 1 and 2 are sent at least once every 2 seconds, and TYPE 3 and 4 are sent at least every second.

The broadcast information is filled by the BSC. Most of the parameters pertain to system configuration, and as such are set by the OSS which indicates them to the BSC as part of the general configuration

Occurrence (modulo 8)	possible messages: SYSTEM INFORMATION TYPE
0	1 if sent at all, or 2, 3, 4, or 2bis
1	2
2	3
3	4
4	1, 2, 3, 4 or 2bis
5	1, 2, 3, 4 or 2bis
6	3
7	4

Table 6.9 – Scheduling of BCCH messages

The four types of BCCH messages are broadcast using a period of 8 occurrences, corresponding to a duration of about 2 seconds.

management. Some parameters must be managed dynamically, and their values may change as the result of local observation. This is the case mainly for the RACH control parameters. As far as the access class is concerned, it can be controlled by BSC only, by OSS only or by both; this is a choice of implementation.

SPECIFICATIONS REFERENCE

The concept of radio resource management as a specific area is introduced basically in **TS GSM 04.07** and in **TS GSM 04.08** (the radio interface application protocols specification), at the start of section 3.

Functional descriptions can be found on some topics in the 03 series. **TS GSM 03.09** deals with the handover function (almost only from the MSC point of view), and **TS GSM 03.13** deals with discontinuous reception.

Power control, measurement reporting and handover preparation are described in detail in **TS GSM 05.08**. Timing advance and synchronisation aspects are dealt with in **TS GSM 05.10**, where the different kinds of handover are described.

TS GSM 04.04 describes the contents of the L1-header in the SACCH messages.

The bulk of the specifications is in fact in the interface specifications, that is to say in:

- **TS GSM 04.08**, section 3, for the RIL3-RR protocol (a part of section 4 is also relevant, since call re-establishment is dealt with in TS GSM 04.08 as part of the mobility management), section 9.5.1 for the description of the messages, and sections 10.5.1 and 10.5.2 for the coding of their information elements;
- **TS GSM 08.58** for the RSM protocol;
- **TS GSM 08.08** for the BSSMAP protocol; and
- **TS GSM 09.02** for the MAP/E protocol. The relevant sections are mainly 5.5 (which deals with handover), plus bits in section 5.7 (operation and maintenance). For those interested in the MAP/VLR interface (the "B" interface, not addressed in this chapter), section 5.15 (paging and search procedures) is also relevant.